

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 13:08:05 ; Search time 123 Seconds
(without alignments)
6501.498 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCUTUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfilesi.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 36 | 2.5 | 2255 | 3 | US-08-871-572B-3 |
| 2 | 35 | 2.4 | 1639 | 3 | Sequence 3, Appli |
| 3 | 35 | 2.4 | 1639 | 3 | Sequence 5, Appli |
| 4 | 34 | 2.4 | 144 | 1 | Sequence 26, Appl |
| 5 | 34 | 2.4 | 348 | 4 | Sequence 13740, A |
| 6 | 34 | 2.4 | 388 | 4 | Sequence 18573, A |
| 7 | 34 | 2.4 | 396 | 4 | Sequence 10, Appl |
| 8 | 34 | 2.4 | 396 | 4 | Sequence 10, Appl |
| 9 | 34 | 2.4 | 397 | 4 | Sequence 18571, A |
| 10 | 34 | 2.4 | 769 | 4 | Sequence 5, Appli |
| 11 | 34 | 2.4 | 1191 | 3 | Sequence 13, Appl |
| 12 | 34 | 2.4 | 1191 | 3 | Sequence 19, Appl |
| 13 | 34 | 2.4 | 1283 | 1 | Sequence 19, Appl |
| 14 | 34 | 2.4 | 1518 | 3 | Sequence 191, App |
| 15 | 34 | 2.4 | 1582 | 3 | Sequence 10, Appl |
| 16 | 34 | 2.4 | 1582 | 3 | Sequence 12, Appl |
| 17 | 34 | 2.4 | 2043 | 4 | Sequence 171, App |
| 18 | 34 | 2.4 | 2103 | 4 | Sequence 40, Appl |
| 19 | 34 | 2.4 | 2218 | 2 | Sequence 4, Appli |
| 20 | 34 | 2.4 | 2218 | 3 | Sequence 31, Appl |
| 21 | 34 | 2.4 | 2780 | 4 | Sequence 87, Appl |
| 22 | 34 | 2.4 | 3244 | 3 | Sequence 4, Appli |
| 23 | 34 | 2.4 | 3437 | 3 | Sequence 17, Appl |
| 24 | 34 | 2.4 | 3437 | 4 | Sequence 17, Appl |
| 25 | 34 | 2.4 | 3581 | 2 | Sequence 1, Appli |
| 26 | 34 | 2.4 | 4434 | 4 | Sequence 861, App |
| 27 | 33 | 2.3 | 229 | 4 | Sequence 10914, A |

| | | | | | | |
|----|----|-----|------|---|---------------------|-------------------|
| 28 | 33 | 2.3 | 273 | 4 | US-09-621-976-11078 | Sequence 11078, A |
| 29 | 33 | 2.3 | 375 | 3 | US-08-946-026-23 | Sequence 23, Appl |
| 30 | 33 | 2.3 | 724 | 3 | US-09-020-956-38 | Sequence 38, Appl |
| 31 | 33 | 2.3 | 724 | 3 | US-09-030-607-38 | Sequence 38, Appl |
| 32 | 33 | 2.3 | 724 | 4 | US-09-439-313-38 | Sequence 38, Appl |
| 33 | 33 | 2.3 | 724 | 4 | US-09-352-616A-38 | Sequence 38, Appl |
| 34 | 33 | 2.3 | 724 | 4 | US-09-232-149A-38 | Sequence 38, Appl |
| 35 | 33 | 2.3 | 724 | 4 | US-09-159-812-38 | Sequence 38, Appl |
| 36 | 33 | 2.3 | 724 | 4 | US-09-636-215-38 | Sequence 38, Appl |
| 37 | 33 | 2.3 | 724 | 4 | US-09-685-166A-38 | Sequence 38, Appl |
| 38 | 33 | 2.3 | 724 | 4 | US-09-115-453-38 | Sequence 38, Appl |
| 39 | 33 | 2.3 | 724 | 4 | US-09-688-489-38 | Sequence 38, Appl |
| 40 | 33 | 2.3 | 911 | 2 | US-08-924-759-9 | Sequence 9, Appli |
| 41 | 33 | 2.3 | 911 | 3 | US-09-248-335-9 | Sequence 9, Appli |
| 42 | 33 | 2.3 | 1052 | 4 | US-09-489-847-23 | Sequence 23, Appl |
| 43 | 33 | 2.3 | 1223 | 3 | US-09-154-874-4 | Sequence 4, Appli |
| 44 | 33 | 2.3 | 1223 | 4 | US-08-931-668-4 | Sequence 4, Appli |
| 45 | 33 | 2.3 | 1223 | 4 | US-09-468-175-4 | Sequence 4, Appli |

ALIGNMENTS

RESULT 1
US-08-871-572B-3
; Sequence 3, Application US/08871572B
; Patent No. 6287853
; GENERAL INFORMATION:
; APPLICANT: Pestka, Sidney
; APPLICANT: Kotenko, Serguei
; APPLICANT: Soh, Jaemog
; APPLICANT: Donnelly, Robert
; APPLICANT: Mariano, Thomas
; APPLICANT: Cook, Jeffrey
; APPLICANT: Emmanuel, Stuart
; APPLICANT: Schwartz, Barbara
; TITLE OF INVENTION: Accessory Factor for Interferon Gamma
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard R. Muccino
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/871,572B
FILING DATE: 9-JUNE-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Muccino, Richard R.
REGISTRATION NUMBER: 32,538
REFERENCE/DOCKET NUMBER: UMD1-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 273-4988
TELEFAX: (908) 273-4679
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2255 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-871-572B-3

```
Query Match      2.5%; Score 36; DB 3; Length 2255;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1406 TTTTAAAAA...AAAAAAAAAAAAAAAAAAAA 1441
DB 2209 TTTTAAAAA...AAAAAAAAAAAAAAAAAAAA 2244

RESULT 2
US-09-362-473-5
; Sequence 5, Application US/09362473
; Patent No. 6218169
; GENERAL INFORMATION:
; APPLICANT: Catoon, Edgar B.
; APPLICANT: Catoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Morgante, Michele
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Hitz, William D.
; APPLICANT: Kinney, Anthony J.
; TITLE OF INVENTION: Aromatic Amino Acid Catabolism Enzymes
; FILE REFERENCE: BB-1197
; CURRENT APPLICATION NUMBER: US/09/362,473
; EARLIER FILING DATE: 1999-07-28
; EARLIER APPLICATION NUMBER: 60/094,783
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1639
; TYPE: DNA
; ORGANISM: Glycine max
US-09-362-473-5

Query Match      2.4%; Score 35; DB 3; Length 1639;
Best Local Similarity 100.0%; Pred. No. 7e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1407 GTTTTAAAAA...AAAAAAAAAAAAAAAAAAAA 1441
DB 1591 GTTTTAAAAA...AAAAAAAAAAAAAAAAAAAA 1625

RESULT 3
US-08-702-344-26
; Sequence 26, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; FILE REFERENCE: ENCODING THEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,344

us-09-936-456-1.oli.rni

FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 144 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-702-344-26

Query Match      2.4%; Score 34; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA...AAAAAAAAAAAAAAAAAAAA 1441
DB 17 TTTTAAAAA...AAAAAAAAAAAAAAAAAAAA 50

RESULT 4
US-09-621-976-13740
; Sequence 13740, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13740
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-13740

Query Match      2.4%; Score 34; DB 4; Length 348;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA...AAAAAAAAAAAAAAAAAAAA 1441
DB 314 TTTTAAAAA...AAAAAAAAAAAAAAAAAAAA 347

RESULT 5
US-09-621-976-18573
; Sequence 18573, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18573
; LENGTH: 388
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-621-976-18573

Query Match 2.4%; Score 34; DB 4; Length 388;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1408 TTTTAAAAA
Db 343 TTTTAAAAA

RESULT 6

US-09-640-173-10/c

; Sequence 10, Application US/09640173
; Patent No. 6613515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
; FILE REFERENCE: 210121.484C2
; CURRENT APPLICATION NUMBER: US/09/640,173
; CURRENT FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G

US-09-640-173-10

Query Match 2.4%; Score 34; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1408 TTTTAAAAA
Db 110 TTTTAAAAA

RESULT 7

US-09-713-550-10/c

; Sequence 10, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.484C4
; CURRENT APPLICATION NUMBER: US/09/713,550
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G

US-09-713-550-10

Query Match 2.4%; Score 34; DB 4; Length 396;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1408 TTTTAAAAA

Db 110 TTTTAAAAA

RESULT 8

US-09-621-976-18571
; Sequence 18571, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18571
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-18571

Query Match 2.4%; Score 34; DB 4; Length 397;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1408 TTTTAAAAA
Db 343 TTTTAAAAA

RESULT 9

US-08-567-882-5

; Sequence 5, Application US/08567882
; Patent No. 6512103
; GENERAL INFORMATION:
; APPLICANT: Dairaghi, Daniel J.
; APPLICANT: Hara, Takahiko
; APPLICANT: Miyajima, Atsushi
; APPLICANT: Schall, Thomas J.
; APPLICANT: Wang, Wei
; APPLICANT: Yoshimura, Akihiko
; TITLE OF INVENTION: MAMMALIAN CHEMOKINE REAGENTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/567,882
; FILING DATE: 08-DEC-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0506
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-852-9196
; TELEFAX: 415-496-1200
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 769 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-567-882-5

Query Match 2.4%; Score 34; DB 4; Length 769;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA 1441
DB 523 TTTTAAAAA 556

RESULT 10
US-09-282-305-13
; Sequence 13, Application US/09282305
; Patent No. 6287843
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/282,305
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/080,563
; PRIOR FILING DATE: 1998-04-03
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)..(971)
US-09-282-305-13

Query Match 2.4%; Score 34; DB 3; Length 1191;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA 1441
DB 1158 TTTTAAAAA 1191

RESULT 11
US-09-883-720-13
; Sequence 13, Application US/09883720
; Patent No. 6479629
; GENERAL INFORMATION:
; APPLICANT: Baldwin, Donald A.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Crane, Virginia C.
; TITLE OF INVENTION: Maize Histone Deacetylases And Their Uses
; FILE REFERENCE: 5718-44,
; CURRENT APPLICATION NUMBER: US/09/883,720
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/282,305
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 13
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (63)..(971)
US-09-883-720-13

Query Match 2.4%; Score 34; DB 4; Length 1191;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA 1441
DB 1158 TTTTAAAAA 1191

RESULT 12
US-08-174-467-19
; Sequence 19, Application US/08174467
; Patent No. 5451514
; GENERAL INFORMATION:
; APPLICANT: BOUDET, ALAIN M.
; APPLICANT: INZE, DIRK G.
; APPLICANT: SCHUCH, WOLFGANG W.
; TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/174,467
; FILING DATE: 28-DEC-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,166
; FILING DATE: 27-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KORULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: 95563/PS36321/US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-174-467-19

Query Match 2.4%; Score 34; DB 1; Length 1283;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1408 TTTTAAAAA 1441
DB 1244 TTTTAAAAA 1277

RESULT 13
US-08-452-071-19
; Sequence 19, Application US/08452071
; Patent No. 6066780
; GENERAL INFORMATION:
; APPLICANT: BOUDET, ALAIN M.
; APPLICANT: INZE, DIRK G.
; APPLICANT: SCHUCH, WOLFGANG W.
; TITLE OF INVENTION: MODIFICATION OF LIGNIN SYNTHESIS IN

;; TITLE OF INVENTION: PLANTS
;; NUMBER OF SEQUENCES: 23
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CUSHMAN, DABY & CUSHMAN
;; STREET: 1100 NEW YORK AVENUE, N.W.
;; CITY: WASHINGTON
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20005
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/452,071
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/874,166
;; FILING DATE: 27-APR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: KOKULIS, PAUL N.
;; REGISTRATION NUMBER: 16,773
;; REFERENCE/DOCKET NUMBER: 95563/PS36321/US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-861-3000
;; TELEFAX: 202-822-0944
;; INFORMATION FOR SEQ ID NO: 19:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1283 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
US-08-452-071-19
Query Match 2.4%; Score 34; DB 3; Length 1283;
Best Local Similarity 100.0%; Pred. No. 1.9e-05; Indels 0; Gaps 0;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1408 TTTTAAAAA 1441
DB 1244 TTTTAAAAA 1277
RESULT 14
US-09-614-912-191
; Sequence 191, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Cai, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/170,906
; PRIOR FILING DATE: 1999-12-15

;; PRIOR APPLICATION NUMBER: 60/172,959
;; PRIOR FILING DATE: 1999-12-21
;; PRIOR APPLICATION NUMBER: 60/172,946
;; PRIOR FILING DATE: 1999-12-21
;; NUMBER OF SEQ ID NOS: 204
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 191
;; LENGTH: 1518
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (114)
;; NAME/KEY: unsure
;; LOCATION: (123)
;; NAME/KEY: unsure
;; LOCATION: (138)
US-09-614-912-191
Query Match 2.4%; Score 34; DB 4; Length 1518;
Best Local Similarity 100.0%; Pred. No. 1.9e-05; Indels 0; Gaps 0;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1408 TTTTAAAAA 1441
DB 1434 TTTTAAAAA 1467
RESULT 15
US-08-545-196B-10
; Sequence 10, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELXI, JUDITH
; APPLICANT: MURNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,196B
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FARACI, C. J.
; REGISTRATION NUMBER: 32,350
; REFERENCE/DOCKET NUMBER: 2121-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-545-196B-10
Query Match 2.4%; Score 34; DB 3; Length 1582;
Best Local Similarity 100.0%; Pred. No. 1.9e-05; Indels 0; Gaps 0;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1408 TTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1441
| | | | |
Db 1485 TTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1518
| | | | |

Search completed: June 3, 2004, 16:17:10
Job time : 126 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 10:21:00 ; Search time 638 Seconds
(without alignments)
9595.068 Million cell updates/sec

Title: US-09-936-456-1

Perfect score: 1441
Sequence: 1 aaagtaacggtacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 337863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1441 | 100.0 | 1441 | 3 | AAA94623 Human CAS |
| 2 | 1413 | 98.1 | 1421 | 3 | AAA96505 cDNA enco |
| 3 | 1336 | 92.7 | 1491 | 6 | ADD19228 Human CDN |
| 4 | 1285 | 89.2 | 1474 | 6 | ABL90535 Human pol |
| 5 | 996 | 69.1 | 1460 | 3 | AAC95564 Human sec |
| 6 | 996 | 69.1 | 1460 | 7 | ABZ67235 Human sec |
| 7 | 996 | 69.1 | 1460 | 7 | ABZ73640 Secreted |
| 8 | 996 | 69.1 | 1460 | 9 | ADC20289 Human sec |
| 9 | 425 | 29.5 | 498 | 3 | AAA94624 Human CAS |
| 10 | 405 | 28.1 | 580 | 9 | ADD19187 Human CDN |
| 11 | 405 | 28.1 | 5033 | 7 | ABZ68115 Human sec |
| 12 | 405 | 28.1 | 5033 | 7 | ABZ74587 Secreted |
| 13 | 405 | 28.1 | 5033 | 9 | ADC21005 Human sec |
| 14 | 392 | 27.2 | 2684 | 7 | ADA53627 Human cod |
| 15 | 303 | 21.0 | 406 | 5 | AAS66563 DNA encod |
| 16 | 59 | 4.1 | 512 | 4 | AH36185 Human col |
| 17 | 38 | 2.6 | 427 | 4 | RAI87437 Human pol |
| 18 | 38 | 2.6 | 999 | 6 | ABN74679 Bovine em |
| 19 | 38 | 2.6 | 2165 | 3 | AAC90468 Human unc |
| 20 | 38 | 2.6 | 5642 | 6 | ABL32869 Human imm |
| 21 | 38 | 2.6 | 10480 | 6 | ABL34201 Human imm |
| 22 | 35 | 2.5 | 896 | 4 | AH33168 Human col |
| 23 | 35 | 2.5 | 2253 | 2 | AAQ84698 Human IFN |

| | | | | | | | |
|---|----|----|-----|------|---|-----------|---------------------|
| c | 24 | 35 | 2.4 | 297 | 5 | ABV61920 | ABV61920 Human pro |
| c | 25 | 35 | 2.4 | 364 | 4 | AAI93544 | AAI93544 Human pol |
| c | 26 | 35 | 2.4 | 436 | 5 | ABV35678 | ABV35678 Human pro |
| c | 27 | 35 | 2.4 | 437 | 5 | ABV44483 | ABV44483 Human pro |
| c | 28 | 35 | 2.4 | 450 | 6 | ABZ08656 | ABZ08656 Human leu |
| c | 29 | 35 | 2.4 | 453 | 4 | AAI87364 | AAI87364 Human pol |
| c | 30 | 35 | 2.4 | 462 | 5 | ABV34102 | ABV34102 Human pro |
| c | 31 | 35 | 2.4 | 462 | 5 | ABV42965 | ABV42965 Human pro |
| c | 32 | 35 | 2.4 | 464 | 6 | ABZ08188 | ABZ08188 Human leu |
| c | 33 | 35 | 2.4 | 504 | 5 | ABV56694 | ABV56694 Human pro |
| c | 34 | 35 | 2.4 | 572 | 5 | ABV61258 | ABV61258 Human pro |
| c | 35 | 35 | 2.4 | 629 | 5 | ABV12980 | ABV12980 Human pro |
| c | 36 | 35 | 2.4 | 644 | 4 | AAI23651 | AAI23651 Human bre |
| c | 37 | 35 | 2.4 | 748 | 4 | AAH71460 | AAH71460 Human cer |
| c | 38 | 35 | 2.4 | 830 | 5 | ABV14600 | ABV14600 Human pro |
| c | 39 | 35 | 2.4 | 1034 | 3 | AAZ52527 | AAZ52527 Human sec |
| c | 40 | 35 | 2.4 | 1639 | 5 | AAAP85102 | AAAP85102 Nucleotid |
| c | 41 | 35 | 2.4 | 2857 | 9 | ADD45824 | ADD45824 Human gen |
| c | 42 | 35 | 2.4 | 2857 | 9 | ADE60246 | ADE60246 Human gen |
| c | 43 | 35 | 2.4 | 2857 | 9 | ADE60248 | ADE60248 Human gen |
| c | 44 | 35 | 2.4 | 2857 | 9 | ADE60252 | ADE60252 Human gen |
| c | 45 | 35 | 2.4 | 2857 | 9 | ADE60250 | ADE60250 Human gen |

ALIGNMENTS

RESULT 1
AAA94623
ID AAA94623 standard; DNA; 1441 BP.
XX
AC AAA94623;
XX
DT 11-JAN-2001 (first entry)
XX
DE Human CASB618 coding sequence.
XX
KW Epitope; human; CASB618; chromosome 15; tumour; vaccine; cancer; ovarian; colon; autoimmune disease; HLA_A0201; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 259..1221
FT /tag= a
FT /product= "Human CASB618"
XX
XX WC00053748-A2.
XX
XX PD 14-SEP-2000.
XX
XX PF 09-MAR-2000; 2000WO-EP002048.
XX
XX PR 11-MAR-1999; 99GB-00005607.
XX
XX PR 01-SEP-1999; 99GB-00020590.
XX
XX PA (SMIX) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX PI Bruck CRM, Cassart J, Coche T, Vinals Y De BassolsC;
XX
XX DR WPI; 2000-572268/53.
XX
XX DR P-PSDB; AAB26325.
XX
XX FT New human CASB618 polypeptide, useful as a vaccine for prophylactic and therapeutic treatment of cancers, particularly ovarian or colon cancer, autoimmune diseases and related conditions.
XX
XX PS Claim 13; Page 61; 76pp; English.
XX
XX CC The present sequence is the coding sequence of human CASB618 protein. The gene for human CASB618 is thought to be located on chromosome 15. The protein encoded by the present sequence and epitopes of the CASB618 protein (see AAB26327 to AAB26399) are useful in diagnosing the occurrence

CC of tumour cells and in vaccines for prophylactic and therapeutic
 CC treatment of cancers, particularly ovarian or colon cancer, autoimmune
 CC diseases and related conditions
 XX

SQ Sequence 1441 BP; 289 A; 466 C; 392 G; 294 T; 0 U; 0 Other;

Query Match 100.0%; Score 1441; DB 3; Length 1441;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTACCGCTACAGACAGTGAAGTATGCTTCGCTCGCGCTAGAGAACTCTGTGG 60
 DB 1 AAAGTACCGCTACAGACAGTGAAGTATGCTTCGCTCGCGCTAGAGAACTCTGTGG 60
 QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCCTTAAACGAGAGGTGCA 120
 DB 61 GTACCAACCCAGAGCGTTGAGAGCAGCCCACTCCACGCTTCCTTAAACGAGAGGTGCA 120
 QY 121 GGACTCAGACTTACAGAGCGCTCGCTCGAGCTTGTGCTTGGCTTGGCTTGGCTTGGCT 180
 DB 121 GGACTCAGACTTACAGAGCGCTCGCTCGAGCTTGTGCTTGGCTTGGCTTGGCTTGGCT 180
 QY 181 GGGCTCTCCGCGCTCCAGAGCAGCCCGAGCTTGTGCTTGGCTTGGCTTGGCTTGGCT 240
 DB 181 GGGCTCTCCGCGCTCCAGAGCAGCCCGAGCTTGTGCTTGGCTTGGCTTGGCTTGGCT 240
 QY 241 ACTGGCCGCGCTGAGAGCGCTTGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 300
 DB 241 ACTGGCCGCGCTGAGAGCGCTTGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 300
 QY 301 CGGCATGCGCGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTAC 360
 DB 301 CGGCATGCGCGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTACAGCTTAC 360
 QY 361 GCAGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 DB 361 GCAGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 421 GTGAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB 421 GTGAGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 481 GAATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 DB 481 GAATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 541 GTTACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 DB 541 GTTACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 601 ACCCCAGTGCATCAGCTGAGAGAGCAATGACTACAAAGCAGTTCACCTGGCGCTCTG 660
 DB 601 ACCCCAGTGCATCAGCTGAGAGAGCAATGACTACAAAGCAGTTCACCTGGCGCTCTG 660
 QY 661 AAAGAGATTAACGCGCGAGTACGCGAAGCGAAGCGGCTGCGGAGCCAGTACAC 720
 DB 661 AAAGAGATTAACGCGCGAGTACGCGAAGCGAAGCGGCTGCGGAGCCAGTACAC 720
 QY 721 CTCTACTGCGGAGAGTTCACACGAGTAGCGCTTGGCGCTTGGCGCTTGGCGCTTGG 780
 DB 721 CTCTACTGCGGAGAGTTCACACGAGTAGCGCTTGGCGCTTGGCGCTTGGCGCTTGG 780
 QY 781 CTGGCGGACACTACGCTCGGCGACGCTATGAGTGGCGTTCGCTTTCGCTTCTCTCC 840
 DB 781 CTGGCGGACACTACGCTCGGCGACGCTATGAGTGGCGTTCGCTTTCGCTTCTCTCC 840
 QY 841 AAGTGTGCTCTCCAGCGCGCGCTCTACGAGGCTGCTGCTGCTGCTGCTGCTGCTG 900
 DB 841 AAGTGTGCTCTCCAGCGCGCGCTCTACGAGGCTGCTGCTGCTGCTGCTGCTGCTG 900
 QY 901 GCCTTCGCGCTCTTTCGCGCTCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 960
 DB 901 GCCTTCGCGCTCTTTCGCGCTCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC 960

RESULT 2

AAA96505
 ID AAA96505 standard; cDNA; 1421 BP.

XX AAA96505;

XX 08-FEB-2001 (first entry)

XX cDNA encoding a human transmembrane protein.

XX Human; transmembrane protein; cell proliferation disorder; myeloma;
 KW reproductive disorder; smooth muscle disorder; neurological disorder;
 KW arteriosclerosis; leukaemia; acquired immunodeficiency syndrome; AIDS;
 KW allergy; ovulatory defect; angina; hypertension; stroke; epilepsy;
 KW Alzheimer's disease; Tourette's disorder; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS /tag= a

XX WO200056891-A2.

XX 28-SEP-2000.

XX 22-MAR-2000; 2000WO-US007817.

XX 22-MAR-1999; 99US-012537P.

XX 16-JUN-1999; 99US-013956SP.

XX (INCY-) INCYTE PHARM INC.

XX Yue H, Lal P, Tang YT, Hillman JL, Reddy R, Bandman O;

PI Baughn MR, Lu DAM, Azimzai Y, Yang J;

XX DR WPI; 2000-579485/54.
 XX DR P-ESDB; AAB18992.
 XX PT New human transmembrane proteins are used to treat a disease or condition
 XX PT associated with decreased expression of functional HMP e.g. Tourette's
 XX PS disorder, angina and leukemia.
 XX PS Claim 4; Page 129; 130pp; English.
 XX CC The present sequence encodes a human transmembrane proteins (HMP).
 XX CC Agonists and antagonists of the protein are used to treat a disease or
 XX CC condition associated with overexpression of the protein. Diseases and
 XX CC conditions which can be treated include cell proliferative disorders
 XX CC immunological, reproductive, smooth muscle and neurological disorders
 XX CC e.g. arteriosclerosis, myeloma, leukemia, acquired immunodeficiency
 XX CC syndrome (AIDS), allergies, ovulatory defects, angina, hypertension,
 XX CC stroke, Alzheimer's disease, epilepsy and Tourette's disorder. The
 XX CC polynucleotides may be used to detect and quantify gene expression in
 XX CC biopsied tissues where protein expression may be correlated with disease
 XX CC e.g. to determine absence, presence or excess expression of HMP or to
 XX CC monitor regulation of HMP expression during therapeutic intervention
 XX SQ Sequence 1421 BP; 264 A; 468 C; 395 G; 294 T; 0 U; 0 Other;
 Query Match 98.1%; Score 1413; DB 3; Length 1421;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGTGGCGGCTAGAAAACTCTGTGC 60
 DB 9 AAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGTGGCGGCTAGAAAACTCTGTGC 68
 QY 61 GTACCAACCCAGAGGGTTGAGAGAGAGCCACCTCCAGCTTCTTAAACGAGAGGTGCA 120
 DB 69 GTACCAACCCAGAGGGTTGAGAGAGAGCCACCTCCAGCTTCTTAAACGAGAGGTGCA 128
 QY 121 GGACTCAGACTTACACAGCCCACTGGTCCAGCCTTTAOCGAAAGAGAGCCCAAGGAC 180
 DB 129 GGACTCAGACTTACACAGCCCACTGGTCCAGCCTTTAOCGAAAGAGAGCCCAAGGAC 188
 QY 181 GGGCTCTCCGGTTCAGAGAGCCAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
 DB 189 GGGCTCTCCGGTTCAGAGAGCCAGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 248
 QY 241 ACTCGGCGGGCTGAGAGATGACCTTGTGGAAGGGGACTGCTCTTTTACCCCGAGGCC 300
 DB 249 ACTCGGCGGGCTGAGAGATGACCTTGTGGAAGGGGACTGCTCTTTTACCCCGAGGCC 308
 QY 301 CGCATGCGCGAGGCTTCAAGGTTCACTGGCTCACTGCTCACTGCTCACTGCTCACTGCTCA 360
 DB 309 CGCATGCGCGAGGCTTCAAGGTTCACTGGCTCACTGCTCACTGCTCACTGCTCACTGCTCA 368
 QY 361 CGAGAGAGCTTCTGCTCATCTTGGCGGATCCGTCGCTGCTGCTGCTGCTGCTGCTGCTG 420
 DB 369 CGAGAGAGCTTCTGCTCATCTTGGCGGATCCGTCGCTGCTGCTGCTGCTGCTGCTGCTG 428
 QY 421 GTGAGAGTCTTCTCAGTCTGTTTCATAGGCGAGAAATGTGGCTGTGCACTTCAGTGCA 480
 DB 429 GTGAGAGTCTTCTCAGTCTGTTTCATAGGCGAGAAATGTGGCTGTGCACTTCAGTGCA 488
 QY 481 GAATGTTTGGTGTGAGTGAAC 540
 DB 489 GAATGTTTGGTGTGAGTGAAC 548
 QY 541 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 DB 549 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608
 QY 601 ACCCAGTGATCAGCTGAACAGAGACCAATGATACAGAGAGTCACTGCTGCTGCTGCTGCT 660
 DB 609 ACCCAGTGATCAGCTGAACAGAGACCAATGATGATACAGAGAGTCACTGCTGCTGCTGCTG 668

QY 661 AAGAGAAATTACGCGCGGAGTAGACGAAACGCACTGGAGAAGGGCTGCGGACCCAGTG 720
 DB 669 AAGAGAAATTACGCGCGGAGTAGACGAAACGCACTGGAGAAGGGCTGCGGACCCAGTG 728
 QY 721 CTCTACTGCGGAGAGAGTTTCAACGAGTAGCCCTTGGGCTGTACGACACAGTACAC 780
 DB 729 CTCTACTGCGGAGAGAGTTTCAACGAGTAGCCCTTGGGCTGTACGACACAGTACAC 788
 QY 781 CTGGCGGACACTACGCTCGGCGACGCTATGGTGGCGTTCTGCTTCTGGCTCTCTCC 840
 DB 789 CTGGCGGACACTACGCTCGGCGACGCTATGGTGGCGTTCTGCTTCTGGCTCTCTCC 848
 QY 841 AAGTGTGCTCTTCCAGCGCGGCGGCTCTTACGAGAGGCTTGGCACTGTGACACCGGA 900
 DB 849 AAGTGTGCTCTTCCAGCGCGGCGGCTCTTACGAGAGGCTTGGCACTGTGACACCGGA 908
 QY 901 GCTTTCGCGCTCTTGGGGTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 960
 DB 909 GCTTTCGCGCTCTTGGGGTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 968
 QY 961 CTCGCGCTTAGGCTCTCCGCGCTCACCACTCAGTACGCGCGGCTTCTGGGTCAAGCTG 1020
 DB 969 CTCGCGCTTAGGCTCTCCGCGCTCACCACTCAGTACGCGCGGCTTCTGGGTCAAGCTG 1028
 QY 1021 GCAACCGGCTCTTGGCTCTTCTCGGAGGGGCGCTGGTGGTGGTGGTGGTGGTGGTGGT 1080
 DB 1029 GCAACCGGCTCTTGGCTCTTCTCGGAGGGGCGCTGGTGGTGGTGGTGGTGGTGGTGG 1088
 QY 1081 CCCAGCGCTCTTCCGACCTTCTTGGACCAAGCGCCAAAGGACTGCAGCCAGGAGAGGG 1140
 DB 1089 CCCAGCGCTCTTCCGACCTTCTTGGACCAAGCGCCAAAGGACTGCAGCCAGGAGAGGG 1148
 QY 1141 GGCTCACTCTTATCTCGGCGACCCACTGCAAGCAGCGCGCTCTCCAGAGACTTAAA 1200
 DB 1149 GGCTCACTCTTATCTCGGCGACCCACTGCAAGCAGCGCGCTCTCCAGAGACTTAAA 1208
 QY 1201 TGTATCACCACCTTGTAGGGGGACCCCAATCTGGACTCTTCCCGCGCTTGGGACAT 1260
 DB 1209 TGTATCACCACCTTGTAGGGGGACCCCAATCTGGACTCTTCCCGCGCTTGGGACAT 1268
 QY 1261 CGCAGGCGGGAAGCAGTGGCGGCGAGCTGCGGCGAGGAGCTCCAGGAGGAGCAGTG 1320
 DB 1269 CGCAGGCGGGAAGCAGTGGCGGCGAGCTGCGGCGAGGAGCTCCAGGAGGAGCAGTG 1328
 QY 1321 AGCGTGTGCGGGGAGGCTGCGGAGATCGCAGGACACAGGAGAAAGTCTCTCGGGGCGA 1380
 DB 1329 AGCGTGTGCGGGGAGGCTTGGGAGATCGGAGACATCCGAGGACCAAGGAGAAAGTCTCTCGGGGCGA 1388
 QY 1381 TCTGTAAATAAACCTTTTTTTTCTTTTGTGTTTTT 1413
 DB 1389 TCTGTAAATAAACCTTTTTTTTCTTTTGTGTTTTT 1421
 RESULT 3
 ADD19228
 ID ADD19228 standard; cDNA; 1491 BP.
 XX
 AC ADD19228;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Human cDNA from secreted protein gene 45.
 KW human secreted protein; cytostatic; antibacterial; virucide;
 KW neuroprotective; gynaecological; gastrointestinal; Gen; cardiac;
 KW cardiovascular; Gen; nephrotropic; anti-inflammatory; muscular; Gen;
 KW respiratory; Gen; immunosuppressive; cerebroprotective; vasotropic;
 KW neurotropic; anti-allergic; cancer; bacterial infection; viral infection;
 KW neural disorder; immune system disorder; blood disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; cardiovascular disorder; renal disorder;
 KW inflammatory disorder; proliferative disorder; human; ss; gene.

Homo sapiens.

WO2003052377-A2.

26-JUN-2003.

06-NOV-2002; 2002WO-US035606.

07-NOV-2001; 2001US-0331046P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM;

WPI; 2003-533050/50.

P-PSDB; AB019303.

New isolated nucleic acids encoding signal transduction pathway component polypeptides, useful for diagnosing, treating, and/or preventing disorders, such as cancer, infections, cardiovascular and inflammatory diseases.

Claim 1; SEQ ID NO 55; 554pp; English.

The invention relates to an isolated nucleic acid molecule (cDNA) encoding a human secreted protein, representing one of 85 novel genes. Also included are recombinant vectors, host cells (expressing the protein), the secreted proteins (including their fragments, epitopes and homologues), an isolated antibody that binds specifically to the protein, diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or absence of a mutation in the nucleic acid and diagnosing a condition based on the presence or absence of the mutation), diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or amount of expression of the protein in a biological sample and diagnosing a condition based on the presence or amount of expression of the protein), preventing, treating or ameliorating a medical condition by administering the nucleic acid or protein to a mammalian subject, identifying a binding partner to the protein, the gene corresponding to the cDNA sequence, and identifying an activity in a biological assay (comprising expressing the nucleic acid in a cell, isolating the supernatant, detecting an activity in a biological assay and identifying the protein in the supernatant having the activity). The nucleic acids and proteins display the following activities: Cytostatic, antibacterial, Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen, Immunosuppressive, Cerebroprotective, Vasotrophic, Nootropic, Antiallergic. The methods and compositions of the present invention are useful for diagnosing, treating, preventing and/or prognosticating disorders related to the novel polypeptides, such as cancer, bacterial or viral infections, and neural, immune system, blood, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, inflammatory or proliferative disorders (many examples of these diseases and disorders are given in the specification). The present sequence encodes a novel secreted protein of the invention.

Sequence 1491 BP; 293 A; 485 C; 410 G; 302 T; 0 U; 1 Other;

Query Match 92.7%; Score 1336; DB 9; Length 1491;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1436; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 AAGATAAGCGCTACAGACAGTGAAGTAATTTGGCTGCGCGCTAGAAAACCTCTGCG 60

44 AAGATAAGCGCTACAGACAGTGAAGTAATTTGGCTGCGCGCTAGAAAACCTCTGCG 103

61 GTACCAACCCACAGAGCGTTGAGAGAGCCACCTCCACGCTTCTTAAAGAGAGGTGCA 120

104 GTACCAACCCACAGAGCGTTGAGAGAGCCACCTCCACGCTTCTTAAAGAGAGGTGCA 163

121 GGACTCAGACTTACAGAGCCACTCGGTCCAGCCTTTGACCAAGAGAGCCCAAGGAC 180

164 GGACTCAGACTTACAGAGCCACTCGGTCCAGCCTTTGACCAAGAGAGCGTCAAGGAC 223

QY 781 CTGGCGGACACTAGCCTCGCCACGCTATGGGTGGGGTCTTCTGCTTCTGCTCTCTCC 840
 DB 802 CTGGCGGACACTAGCCTCGCCACGCTATGGGTGGGGTCTTCTGCTTCTGCTCTCTCC 861
 QY 841 AACGTGCTGCTCTTCCACGCGCGCGCTCTACGAGAGGCTTGGCACTCTGACACCGGA 900
 DB 862 AACGTGCTGCTCTTCCACGCGCGCGCTCTACGAGAGGCTTGGCACTCTGACACCGGA 921
 QY 901 GCTTCGCGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCT 960
 DB 922 GCTTCGCGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCT 981
 QY 961 CTCGCGCTAGGCTCTCTCCGCGCTCACCACTCAGTACGCGCGCGCTTCTGGCTCAGCTG 1020
 DB 982 CTCGCGCTAGGCTCTCTCCGCGCTCACCACTCAGTACGCGCGCGCTTCTGGCTCAGCTG 1041
 QY 1021 GCAACCGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCT 1047
 DB 1042 GCAACCGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCTCTTGGGCT 1068

RESULT 7
 ABZ73640
 ID ABZ73640 standard; cDNA; 1460 BP.
 XX
 AC ABZ73640;
 CC
 DT 12-MAY-2003 (first entry)
 XX
 DS Secreted protein-encoding gene 360 cDNA clone HUFCL31, SEQ ID NO:370.
 XX
 KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
 KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 KW drug screening; chromosome identification; chromosome mapping;
 KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
 KW antianaemic; vulnery; chromosome 9p21; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200277013-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 26-MAR-2002; 2002WO-US009370.
 XX
 PR 27-MAR-2001; 2001US-027850P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 XX
 XX WPI; 2003-040578/03.
 XX
 XX P-PSDB; ABR01306.
 XX
 XX New human secreted proteins and nucleic acids; useful for detecting or
 XX treating cancer or other hyperproliferative disorders, autoimmune
 XX disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
 XX
 XX Claim 21; Page 1345; 2474pp; English.
 XX
 XX ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
 XX protein genes, and ABP0947-ABP01363 represent the proteins they encode.
 XX
 XX ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
 XX invention also encompasses antibodies specific for the secreted proteins.
 XX
 XX the use of the secreted proteins in drug screening and recombinant
 XX vectors and host cells comprising a nucleic acid of the invention. The
 XX secreted proteins are thought to be involved in biological activities
 XX associated with cellular signalling, cellular differentiation, cell
 XX migration, pro-hormone activation and neurotransmitter activity. The
 XX secreted proteins, nucleic acids encoding them, antibodies or antibody

fragments specific for the secreted proteins, and modulators of protein
 activity are useful for diagnosing or treating cancers or other
 hyperproliferative disorders. Additionally, the secreted proteins and
 their nucleic acids may also be used in the treatment of autoimmune
 disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
 (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
 wound healing. Nucleic acids of the invention may be used for chromosome
 identification, chromosome mapping, in gene therapy, for identifying
 individuals from minute biological samples, as hybridisation probes, and
 as molecular weight markers. The present sequence represents a human
 secreted protein-encoding cDNA clone of the invention

Query Match 69.1%; Score 996; DB 7; Length 1460;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1046; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AAAGTACGGCTACGACAGTGAAGATAGTTTCGCTCGCGGCTAGAAAACCTCTGCG 60
 DB 22 AAAGTACGGCTACGACAGTGAAGATAGTTTCGCTCGCGGCTAGAAAACCTCTGCG 81
 QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAAACGAGAGGTGCA 120
 DB 82 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAAACGAGAGGTGCA 141
 QY 121 GGAATCAGCTTCAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 DB 142 GGAATCAGCTTCAACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 201
 QY 181 GCGCTCTCCGCTCCAGGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 DB 202 GCGCTCTCCGCTCCAGGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 261
 QY 241 ACTCGCGCGGCTGAGCATGACCTGTGGAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 DB 262 ACTCGCGCGGCTGAGCATGACCTGTGGAACGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
 QY 301 CGGATGCGCGAGGCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 DB 322 CGGATGCGCGAGGCTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381
 QY 361 GCAGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 DB 382 GCAGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441
 QY 421 GTGAGAGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB 442 GTGAGAGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 501
 QY 481 GAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 DB 502 GAATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
 QY 541 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 DB 562 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
 QY 601 ACCCGAGTGCATGCTGAGACGAGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 DB 622 ACCCGAGTGCATGCTGAGACGAGCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 681
 QY 661 AAAGAGAAATTAACGCGGAGTACGCAACGCTGAGAGAGGCTGCGGAGCCGAGT 720
 DB 682 AAAGAGAAATTAACGCGGAGTACGCAACGCTGAGAGAGGCTGCGGAGCCGAGT 741
 QY 721 CTCTACCTGCGGAGAGTTCACCGAGTACGCTTGGGCTGTGACACACAGTACAC 780
 DB 742 CTCTACCTGCGGAGAGTTCACCGAGTACGCTTGGGCTGTGACACACAGTACAC 801
 QY 781 CTGGCGGAGCACTACGCTCTCGGCGAGCTATGGGTGGGCTTCTGCTTCTGCTCTCTCC 840
 DB 802 CTGGCGGAGCACTACGCTCTCGGCGAGCTATGGGTGGGCTTCTGCTTCTGCTCTCTCC 861

CC gastroenteritis). The present DNA sequence encodes a human secreted
CC protein of the invention.
XX
SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;
Query Match 69.1%; Score 996; DB 9; Length 1460;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1046; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGTCGCGGGCTAGAAAACTCTGTCG 60
DB 22 AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGTCGCGGGCTAGAAAACTCTGTCG 81
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGAGAGAGTGCA 120
DB 82 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCCTTAACGAGAGAGTGCA 141
QY 121 GGACTCAGACTTCCAGAGCCACTCGGTCCAGCCTTGTACGCAAGAGAGACCAAGAC 180
DB 142 GGACTCAGACTTCCAGAGCCACTCGGTCCAGCCTTGTACGCAAGAGAGACCAAGAC 201
QY 181 CGGCTCTCCGCGTCCAGAGCGGAGCCAGCTTGCCTGCTGCTGCTGCTGCTGCTGCTG 240
DB 202 CGGCTCTCCGCGTCCAGAGCGGAGCCAGCTTGCCTGCTGCTGCTGCTGCTGCTGCTG 261
QY 241 ACTCGCGCGGCTGCGAGCATGACCTGTGGAACGGCGTACTGCTCTTTTACCCCGAGCC 300
DB 262 ACTCGCGCGGCTGCGAGCATGACCTGTGGAACGGCGTACTGCTCTTTTACCCCGAGCC 321
QY 301 CGGATGCGCGAGGCTTCCAGCTGCTCATGCTATCTAGTCTGCTGCTGCTGCTGCTGCTG 360
DB 322 CGGATGCGCGAGGCTTCCAGCTGCTCATGCTATCTAGTCTGCTGCTGCTGCTGCTGCTG 381
QY 361 GCAGCAAGCTTCTGCTCATCTGCGGGGATCCGTCGCGCACTGCGCTGCTGCTGCTGCTG 420
DB 382 GCAGCAAGCTTCTGCTCATCTGCGGGGATCCGTCGCGCACTGCGCTGCTGCTGCTGCTG 441
QY 421 GTGAGAGTTCTTCTCAGTCTGTTCATAGCGGAGAAATGCTGCTGCTGCTGCTGCTGCTG 480
DB 442 GTGAGAGTTCTTCTCAGTCTGTTCATAGCGGAGAAATGCTGCTGCTGCTGCTGCTGCTG 501
QY 481 GAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
DB 502 GAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 561
QY 541 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 562 GTTACAGCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
QY 601 ACCCCAGTGCATCAGCTGAAACGAGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
DB 622 ACCCCAGTGCATCAGCTGAAACGAGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 681
QY 661 AAAGAGAAATACCGCGCGAGTACCGGAAACGCACTGGAGAGAGGGCTGCGGACCCAGTG 720
DB 682 AAAGAGAAATACCGCGCGAGTACCGGAAACGCACTGGAGAGAGGGCTGCGGACCCAGTG 741
QY 721 CTCTACCTGGCGGAGAGTTCACACGAGTAGCCCTTGGGGCTGTACACAGTACACAC 780
DB 742 CTCTACCTGGCGGAGAGTTCACACGAGTAGCCCTTGGGGCTGTACACAGTACACAC 801
QY 781 CTGGCGGAGACACTACGCTCTCGGCGCACGCTATGGGTGGGCTTCTGCTTCTGGCTCTCTCC 840
DB 802 CTGGCGGAGACACTACGCTCTCGGCGCACGCTATGGGTGGGCTTCTGCTTCTGGCTCTCTCC 861
QY 841 AAAGTGTGCTCTCCACGCGCGGCGCTCTACGAGAGGCTGCGACTGCTGCTGCTGCTGCTG 900
DB 862 AAAGTGTGCTCTCCACGCGCGGCGCTCTACGAGAGGCTGCGACTGCTGCTGCTGCTGCTG 921
QY 901 GCCTTCGCGCTCTCTCGGGTCTTCCGCTTGGCTTCCATCTTAGGTGCTGCTGCTGCTGCTG 960
DB 922 GCCTTCGCGCTCTCTCGGGTCTTCCGCTTGGCTTCCATCTTAGGTGCTGCTGCTGCTGCTG 981

QY 841 AACGTGCTGCTCTCCACGCGGCGGCTCTACGAGAGGCTGCGACTGCTGACCAACCGGA 900
DB 862 AACGTGCTGCTCTCCACGCGGCGGCTCTACGAGAGGCTGCGACTGCTGACCAACCGGA 921
QY 901 GCCTTCGCGCTCTCTCGGGTCTTCCGCTTGGCTTCCATCTTAGGTGCTGCTGCTGCTG 960
DB 922 GCCTTCGCGCTCTCTCGGGTCTTCCGCTTGGCTTCCATCTTAGGTGCTGCTGCTGCTG 981
QY 961 CTCGCGCTAGGCTCTCTCGCGCTCACCACTCAGTACGCGCGCGCTTCTGCGTACGCTG 1020
DB 982 CTCGCGCTAGGCTCTCTCGCGCTCACCACTCAGTACGCGCGCGCTTCTGCGTACGCTG 1041
QY 1021 GCAACGCGGCTCTGCTGCTCTTCTCTC 1047
DB 1042 GCAACGCGGCTCTGCTGCTCTTCTCTC 1068
RESULT 8
ADC20289
ID ADC20289 standard; DNA; 1460 BP.
AC
XX
AC
XX
DT 18-DEC-2003 (first entry)
XX
DE Human secreted protein coding sequence #228.
XX
KW gene therapy; human; secreted protein; haemopoietic disorder;
KW haematological disorder; anaemia; haemophilia; inflammatory disorder;
KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;
KW leukaemia; wound healing; epithelial cell proliferation disorder;
KW immune disorder; autoimmune disorder; asthmatic disorder;
KW cardiovascular disorder; atherosclerosis; myocarditis;
KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;
KW gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.
XX
OS Homo sapiens.
XX
XX WO200292787-A2.
XX
XX 21-NOV-2002.
XX
XX 26-MAR-2002; 2002WO-US009257.
XX
XX 27-MAR-2001; 2001US-0278650P.
PR 12-SEP-2001; 2001US-00950082.
PR 12-SEP-2001; 2001US-00950083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2003-129287/12.
XX
XX New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,
PT preventing or treating hematopoietic or hematologic disorders, e.g.
PT anemia or hemophilia.
XX
XX Claim 1; SEQ ID NO 238; 1512pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC secreted proteins. The DNA and protein sequences of the invention are
CC useful for detecting, preventing, diagnosing, prognosticating, treating
CC or ameliorating; haematopoietic or haematological disorders (e.g. anaemia
CC and haemophilia); inflammatory disorders (e.g. inflammatory bowel disease
CC and Crohn's disease); neoplastic disease (e.g. cancer and leukaemia);
CC wound healing and disorders of epithelial cell proliferation; immune
CC disorders (e.g. autoimmune disorders and asthmatic disorders);
CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC and gastrointestinal disorders (e.g. duodenal ulcers and

CC presence or amount of expression of the protein in a biological sample
 CC and diagnosing a condition based on the presence or amount of expression
 CC of the protein), preventing, treating or ameliorating a medical condition
 CC by administering the nucleic acid or protein to a mammalian subject,
 CC identifying a binding partner to the protein, the gene corresponding to
 CC the cDNA sequence, and identifying an activity in a biological assay
 CC (comprising expressing the nucleic acid in a cell, isolating the
 CC supernatant, detecting an activity in a biological assay and identifying
 CC the protein in the supernatant having the activity). The nucleic acids
 CC and proteins display the following activities: Cytostatic, antibacterial,
 CC Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen,
 CC Cardiac, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-
 CC Gen, Respiratory-Gen, Immunosuppressive, Cerebroprotective, Vasotropic,
 CC Nootropic, Antiallergic. The methods and compositions of the present
 CC invention are useful for diagnosing, treating, preventing and/or
 CC prognosticating disorders related to the novel polypeptides, such as
 CC cancer, bacterial or viral infections, and neural, immune system, blood,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal, inflammatory or proliferative disorders (many examples of these
 CC diseases and disorders are given in the specification). The present
 CC sequence encodes a novel secreted protein of the invention.

XX Sequence 580 BP; 115 A; 198 C; 156 G; 111 T; 0 U; 0 Other;
 Query Match 28.1%; Score 405; DB 9; Length 580;
 Best Local Similarity 100.0%; Pred. No. 8.6e-146;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCCTCGCGGCTAGAAAACCTCTGCG 60
 DB 13 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCCTCGCGGCTAGAAAACCTCTGCG 72
 QY 61 GTACCAACCCAGAGCGTTGAGACAGCCCACTCCAGCTTCCTTAACGAGAGGTGCA 120
 DB 73 GTACCAACCCAGAGCGTTGAGACAGCCCACTCCAGCTTCCTTAACGAGAGGTGCA 132
 QY 121 GGACTCAGACTTACAGCCCACTCGGTCCAGCCTTGTACGAAAGAGAGCCGAGGAC 180
 DB 133 GGACTCAGACTTACAGCCCACTCGGTCCAGCCTTGTACGAAAGAGAGCCGAGGAC 192
 QY 181 GCGCTCTCCCGGTTCAGAGCAGCCCACTCGGTCCAGCCTTGTACGAAAGAGAGCCGAGGAC 240
 DB 193 GCGCTCTCCCGGTTCAGAGCAGCCCACTCGGTCCAGCCTTGTACGAAAGAGAGCCGAGGAC 252
 QY 241 ACTCGGCGCGGTGAGCAGATGACCTGTGGAAGCGGCTATGCTTTTACCCCGAGGCC 300
 DB 253 ACTCGGCGCGGTGAGCAGATGACCTGTGGAAGCGGCTATGCTTTTACCCCGAGGCC 312
 QY 301 CGGCATGCGCGAGGCTTCAGCGTTCACCTGCTATCTATCTAGTGTATTTTGGCTCTA 360
 DB 313 CGGCATGCGCGAGGCTTCAGCGTTCACCTGCTATCTATCTAGTGTATTTTGGCTCTA 372
 QY 361 GCAGCAAGCTTCTGCTCATCTTTGCGGGGATCCGTGGCCACTCG 405
 DB 373 GCAGCAAGCTTCTGCTCATCTTTGCGGGGATCCGTGGCCACTCG 417

RESULT 11
 ABZ68115
 ID ABZ68115 standard; DNA; 5033 BP.

XX AC ABZ68115;

XX DT 26-MAR-2003 (first entry)

XX DE Human secreted protein encoding genomic DNA SEQ ID NO 1638.

XX KW Human; secreted protein; neurotropic; neuroprotective; cytostatic;
 KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
 KW vulnary; antibacterial; antiparkinsonian; antiskilling; antianemic;
 KW antiarthritic; cancer; antineumatic; hepatotropic; cerebroprotective;
 KW antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant;
 KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;

KW cardiovascular disorder; neurological disease; nephrotropic;
 KW gene therapy; gene; ds.
 OS Homo sapiens.
 XX WO200277186-A2.
 PN 03-OCT-2002.
 PD 26-MAR-2002; 2002WO-US009188.
 XX 27-MAR-2001; 2001US-0278650P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM;
 WIPI; 2003-040583/03.
 PT New human secreted proteins encoded by genes contained in cDNA clones
 PT (e.g. HGCAC19), useful for preventing, treating or diagnosing e.g. AIDS,
 PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
 PT West Nile fever.
 PS Disclosure; Page 2263-2264; 2423pp; English.

CC The invention relates to novel human genes (ABZ68891-ABZ68209) and the
 CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
 CC treating or ameliorating medical conditions e.g. by protein or gene
 CC therapy. The genes are isolated from a range of human tissues disclosed
 CC in the specification. The nucleic acids, proteins, antibodies and
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
 CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
 CC myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g.
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections

XX Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match 28.1%; Score 405; DB 7; Length 5033;
 Best Local Similarity 100.0%; Pred. No. 6e-146;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCCTCGCGGCTAGAAAACCTCTGCG 60
 DB 13 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTCCTCGCGGCTAGAAAACCTCTGCG 72
 QY 61 GTACCAACCCAGAGCGTTGAGACAGCCCACTCCAGCTTCCTTAACGAGAGGTGCA 120
 DB 73 GTACCAACCCAGAGCGTTGAGACAGCCCACTCCAGCTTCCTTAACGAGAGGTGCA 132
 QY 121 GGACTCAGACTTACAGCCCACTCGGTCCAGCCTTGTACGAAAGAGAGCCGAGGAC 180
 DB 133 GGACTCAGACTTACAGCCCACTCGGTCCAGCCTTGTACGAAAGAGAGCCGAGGAC 192
 QY 181 GCGCTCTCCCGGTTCAGAGCAGCCCACTCGGTTCGCTTGTGCTTGCCTGCGGCTGCGGACG 240
 DB 193 GCGCTCTCCCGGTTCAGAGCAGCCCACTCGGTTCGCTTGTGCTTGCCTGCGGCTGCGGACG 252
 QY 241 ACTCGGCGCGGTGAGCAGATGACCTGTGGAAGCGGCTATGCTTTTACCCCGAGGCC 300
 DB 253 ACTCGGCGCGGTGAGCAGATGACCTGTGGAAGCGGCTATGCTTTTACCCCGAGGCC 312
 QY 301 CGGCATGCGCGAGGCTTCAGCGTTCACCTGCTATCTATCTAGTGTATTTTGGCTCTA 360
 DB 313 CGGCATGCGCGAGGCTTCAGCGTTCACCTGCTATCTATCTAGTGTATTTTGGCTCTA 372

QY 361 GCACGAGCTTCCTGCTCATCTTCCCGGGGATCCGTGGCCACTCG 405
 DB 373 GCACGAGCTTCCTGCTCATCTTCCCGGGGATCCGTGGCCACTCG 417

RESULT 12

ABZ74587

ID ABZ74587 standard; DNA; 5033 BP.

XX AC ABZ74587;

XX DT 12-MAY-2003 (first entry)

XX DE Secreted protein gene 360 genomic fragment HUPCL31, SEQ ID NO:1734.

XX KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
 XX KW autoimmune disorder; inflammation; angiotensin diseases; AIDS;
 XX KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
 XX KW drug screening; chromosome identification; chromosome mapping;
 XX KW cytoskeletal; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
 XX KW antianaemic; vulnery; chromosome 9p21; gene; ds.

XX OS Homo sapiens.

XX PN WO200277013-A2.

XX PD 03-OCT-2002.

XX PF 26-MAR-2002; 2002WO-US009370.

XX PR 27-MAR-2001; 2001US-0278650P.

XX PR 12-SEP-2001; 2001US-00950082.

XX PR 12-SEP-2001; 2001US-00950083.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR WPI; 2003-040578/03.

XX PT New human secreted proteins and nucleic acids, useful for detecting or

XX PT treating cancer or other hyperproliferative disorders, autoimmune

XX PT disorders, inflammatory disorders, HIV disease, hepatitis or anaemia.

XX PS Disclosure; Page 2315-2316; 2474pp; English.

XX CC ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted

XX CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.

XX CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The

XX CC invention also encompasses antibodies specific for the secreted proteins,

XX CC the use of the secreted proteins in drug screening and recombinant

XX CC vectors and host cells comprising a nucleic acid of the invention. The

XX CC secreted proteins are thought to be involved in biological activities

XX CC associated with cellular signalling, cellular differentiation, cell

XX CC migration, prohormone activation and neurotransmitter activity. The

XX CC secreted proteins, nucleic acids encoding them, antibodies or antibody

XX CC fragments specific for the secreted proteins, and modulators of protein

XX CC activity are useful for diagnosing or treating cancers or other

XX CC hyperproliferative disorders. Additionally, the secreted proteins and

XX CC their nucleic acids may also be used in the treatment of autoimmune

XX CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS

XX CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote

XX CC wound healing. Nucleic acids of the invention may be used for chromosome

XX CC identification, chromosome mapping, in gene therapy, for identifying

XX CC individuals from minute biological samples, as hybridisation probes, and

XX CC as molecular weight markers. The present sequence represents a human

XX CC secreted protein genomic fragment referred to in the disclosure of the

XX CC invention

XX SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match 28.1%; Score 405; DB 7; Length 5033;

Best Local Similarity 100.0%; Pred. No. 6e-146;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAAGTAAAGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGCTAGAAAAAATCTGTGTG 60
 DB 13 AAAGTAAAGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGCTAGAAAAAATCTGTGTG 72
 QY 61 GTACCAACCCAGAGCGTTGAGAGAGCGCCACCTCCACCGCTTCTTAACGAGAGGTGCA 120
 DB 73 GTACCAACCCAGAGCGTTGAGAGAGCGCCACCTCCACCGCTTCTTAACGAGAGGTGCA 132
 QY 121 GGACTCAGACTTCACCAAGCCCACTCGGTCCAGCCTTGTACGCAAGAGAGCGCAAGGAC 180
 DB 133 GGACTCAGACTTCACCAAGCCCACTCGGTCCAGCCTTGTACGCAAGAGAGCGCAAGGAC 192
 QY 181 GCGCTCTCCCGCTCCAGGCGAGCCAGCTTGTGCTTGCCTGCGCTCGCGCTGCTGCAC 240
 DB 193 GCGCTCTCCCGCTCCAGGCGAGCCAGCTTGTGCTTGCCTGCGCTCGCGCTGCTGCAC 252
 QY 241 ACTCGGCGCGCTGTCAGCATGACCTTGTGGAACGGGTACTGCTTTTACCCCGAGCCC 300
 DB 253 ACTCGGCGCGCTGTCAGCATGACCTTGTGGAACGGGTACTGCTTTTACCCCGAGCCC 312
 QY 301 CGGCTATGCCGCGGCTTCAGCGTTCCACCTGCTCATCTTATCTAGTGTGTTGGCTCTA 360
 DB 313 CGGCTATGCCGCGGCTTCAGCGTTCCAGCTGCTCATCTTATCTAGTGTGTTGGCTCTA 372
 QY 361 GCAGCAAGCTTCCTGCTCATCTTTCGCGGGGATCCGTGGCCACTCG 405
 DB 373 GCAGCAAGCTTCCTGCTCATCTTTCGCGGGGATCCGTGGCCACTCG 417

RESULT 13

ADC21005

ID ADC21005 standard; DNA; 5033 BP.

XX AC ADC21005;

XX DT 18-DEC-2003 (first entry)

XX DE Human secreted protein-related DNA sequence #423.

XX KW gene therapy; human; secreted protein; haemopoietic disorder;

XX KW haematological disorder; anaemia; haemophilia; inflammatory disorder;

XX KW inflammatory bowel disease; Crohn's disease; neoplastic disease; cancer;

XX KW leukaemia; wound healing; epithelial cell proliferation disorder;

XX KW immune disorder; autoimmune disorder; asthmatic disorder;

XX KW cardiovascular disorder; atherosclerosis; myocarditis;

XX KW infectious disease; HIV; AIDS; endocrine disorder; diabetes;

XX KW gastrointestinal disorder; duodenal ulcer; gastroenteritis; gene; ds.

XX OS Homo sapiens.

XX PN WO200292787-A2.

XX PD 21-NOV-2002.

XX PF 26-MAR-2002; 2002WO-US009257.

XX PR 27-MAR-2001; 2001US-0278650P.

XX PR 12-SEP-2001; 2001US-00950082.

XX PR 12-SEP-2001; 2001US-00950083.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX DR WPI; 2003-129287/12.

XX PT New human secreted proteins and nucleic acid molecules, useful for

XX PT preparing a diagnostic or pharmaceutical composition for diagnosing,

XX PT preventing or treating hematopoietic or hematologic disorders, e.g.

XX PT anemia or hemophilia.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 08:29:22 ; Search time 5809 Seconds
(without alignments)
10751.818 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaaa 1441

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|-----------|---------------------|
| 1 | 1441 | 100.0 | 1441 | 6 | AX035346 | AX035346 Sequence |
| 2 | 1029 | 71.4 | 1740 | 9 | HSM806241 | BX537581 Homo sapi |
| 3 | 711.4 | 49.4 | 1354 | 10 | BC031111 | BC031111 Mus muscu |
| 4 | 626.2 | 43.5 | 5851 | 4 | AF547266 | AF547266 Sus scrofa |
| 5 | 465.4 | 32.3 | 498 | 6 | AX035348 | AX035348 Sequence |
| 6 | 405.4 | 28.1 | 166937 | 2 | AC012255 | AC012255 Homo sapi |
| 7 | 405.4 | 28.1 | 181312 | 9 | AC091117 | AC091117 Homo sapi |
| 8 | 403.8 | 28.0 | 156534 | 2 | AC009700 | AC009700 Homo sapi |
| 9 | 403.8 | 28.0 | 171444 | 2 | AC087790 | AC087790 Homo sapi |
| 10 | 396 | 27.5 | 2684 | 6 | AX714511 | AX714511 Sequence |
| 11 | 396 | 27.5 | 2684 | 9 | AX056896 | AX056896 Homo sapi |
| 12 | 396 | 27.5 | 156534 | 2 | AC009700 | AC009700 Homo sapi |
| 13 | 356.4 | 24.7 | 1636 | 6 | AX045919 | AX045919 Sequence |
| 14 | 356.4 | 24.7 | 1923 | 9 | BC029819 | BC029819 Homo sapi |
| 15 | 348.8 | 24.2 | 1521 | 10 | BC019755 | BC019755 Mus muscu |
| 16 | 299.4 | 20.8 | 2439 | 9 | AK027187 | AK027187 Homo sapi |
| 17 | 245.8 | 17.1 | 1706 | 9 | EC020841 | EC020841 Homo sapi |
| 18 | 245.8 | 17.1 | 2684 | 6 | AX714511 | AX714511 Sequence |
| 19 | 245.8 | 17.1 | 2684 | 9 | AX056896 | AX056896 Homo sapi |
| 20 | 217 | 15.1 | 217 | 9 | HS183H12F | Z57552 H.sapiens C |
| 21 | 199.8 | 13.9 | 218 | 9 | HS183A12R | Z59954 H.sapiens C |
| 22 | 167.8 | 11.6 | 229583 | 10 | AL844566 | AL844566 Mouse DNA |
| 23 | 166 | 11.5 | 1923 | 9 | BC029819 | BC029819 Homo sapi |
| 24 | 165.8 | 11.5 | 231868 | 2 | AC118124 | AC118124 Rattus no |
| 25 | 159.4 | 11.1 | 166937 | 2 | AC012255 | AC012255 Homo sapi |
| 26 | 159.4 | 11.1 | 171444 | 2 | AC087790 | AC087790 Homo sapi |
| 27 | 159.4 | 11.1 | 181312 | 9 | AC091117 | AC091117 Homo sapi |
| 28 | 154.2 | 10.7 | 231868 | 2 | AC118124 | AC118124 Rattus no |
| 29 | 147 | 10.2 | 229583 | 10 | AL844566 | AL844566 Mouse DNA |
| 30 | 123.2 | 8.5 | 110250 | 2 | AC138220 | AC138220 Mus muscu |
| 31 | 121.6 | 8.4 | 55061 | 2 | AC091597 | AC091597 Mus muscu |
| 32 | 101 | 7.0 | 110250 | 2 | AC138220 | AC138220 Mus muscu |
| 33 | 65.6 | 4.6 | 2207 | 3 | AY061346 | AY061346 Drosophila |
| 34 | 63.4 | 4.4 | 125020 | 9 | AF429315 | AF429315 Homo sapi |
| 35 | 61 | 4.2 | 7218 | 6 | I66494 | I66494 Sequence 14 |
| 36 | 60.6 | 4.2 | 2000 | 6 | AX655393 | AX655393 Sequence |
| 37 | 59.6 | 4.1 | 125020 | 9 | AF429315 | AF429315 Homo sapi |
| 38 | 56.8 | 3.9 | 63689 | 2 | AC024268 | AC024268 Homo sapi |
| 39 | 55.8 | 3.9 | 4615 | 1 | STWHIB12X | L22864 Streptomyces |
| 40 | 55.2 | 3.8 | 151173 | 8 | AC135257 | AC135257 Genomic s |
| 41 | 53 | 3.7 | 2690 | 8 | AY332473 | AY332473 Oryza sat |
| 42 | 53 | 3.7 | 169162 | 8 | AP004267 | AP004267 Oryza sat |
| 43 | 52 | 3.6 | 2000 | 6 | AX655393 | AX655393 Sequence |
| 44 | 52 | 3.6 | 117787 | 2 | AC104271 | AC104271 Oryza sat |
| 45 | 51.8 | 3.6 | 168866 | 8 | OSJN00023 | AL606588 Oryza sat |

ALIGNMENTS

RESULT 1
AX035346
LOCUS AX035346
DEFINITION Sequence 1 from Patent WO053748.
ACCESSION AX035346
VERSION AX035346.1 GI:11191064
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.
TITLE Novel compounds
JOURNAL Patent: WO 0053748-A 1 14-SEP-2000;

PAT 15-NOV-2000

BRUCK CLAUDELINE ELVIRE MARIE (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ;
VINALS Y DE BASSOLS CARLOTA (BE) ; COCHE THIERRY (BE) ; CASSART
JEAN POL (BE)

FEATURES
source 1. 1441
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1441; DB 6; Length 1441;
Best Local Similarity 100.0%; Pred. No. 3.3e-294;
Matches 1441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGTAAGCGCTACAGACAGTGAAGAAATAGTTTGGCTCGCGCGGTAGAAAACCTCTCG 60
Db 1 AAAGTAAGCGCTACAGACAGTGAAGAAATAGTTTGGCTCGCGCGGTAGAAAACCTCTCG 60
Qy 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAAACGGAGAGTGA 120
Db 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAAACGGAGAGTGA 120
Qy 121 GGACTCAGACTTACAGACCCCACTCGTCCAGCGCTTGTACGAAAGAGAGCCAGGAC 180
Db 121 GGACTCAGACTTACAGACCCCACTCGTCCAGCGCTTGTACGAAAGAGAGCCAGGAC 180
Qy 181 GCGCTCTCCCGCTCCAGCGAGCCCGACCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 240
Db 181 GCGCTCTCCCGCTCCAGCGAGCCCGACCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 240
Qy 241 ACTCGCGCGCGGTGAGCATGACCTGTGGAACCGCGTACTGCTTTTACCCCGAGGCC 300
Db 241 ACTCGCGCGCGGTGAGCATGACCTGTGGAACCGCGTACTGCTTTTACCCCGAGGCC 300
Qy 301 CGGCATCGCGAGCGTTCAGCGTTCACCTGCTATCTGCTTATCTAGTCTTTTGGCTCTA 360
Db 301 CGGCATCGCGAGCGTTCAGCGTTCACCTGCTATCTGCTTATCTAGTCTTTTGGCTCTA 360
Qy 361 GCAGCAAGCTTCTGCTCTCATCTTTCGCGGGATCCGTCGCGCACTCGCGCTGTTGGTTG 420
Db 361 GCAGCAAGCTTCTGCTCTCATCTTTCGCGGGATCCGTCGCGCACTCGCGCTGTTGGTTG 420
Qy 421 GTGAGAGTCTTCTCAGTCTGTTCATAGCGGAGAAATGTGGCTGTCACCTTCAGTGCA 480
Db 421 GTGAGAGTCTTCTCAGTCTGTTCATAGCGGAGAAATGTGGCTGTCACCTTCAGTGCA 480
Qy 481 GAATGGTTCGTCGGTACAGTGAACCAACACATCTCTCAAAAGCCCTTCAGCGAGCGGC 540
Db 481 GAATGGTTCGTCGGTACAGTGAACCAACACATCTCTCAAAAGCCCTTCAGCGAGCGGC 540
Qy 541 GTTACAGCCCGTGTGCTGCTCGTGGCCCTGGAGGGCAATTAATTACACTCAGAGG 600
Db 541 GTTACAGCCCGTGTGCTGCTCGTGGCCCTGGAGGGCAATTAATTACACTCAGAGG 600
Qy 601 ACCCAGTGCATCAGTGAACAGACCACTTGAACAGAGAGTTCACCTGGCGTCTG 660
Db 601 ACCCAGTGCATCAGTGAACAGACCACTTGAACAGAGAGTTCACCTGGCGTCTG 660
Qy 661 AAAGAGAAATTACCGCGGAGTACGGAACGCACTGGAGAGGGGTGCGGACCCAGTG 720
Db 661 AAAGAGAAATTACCGCGGAGTACGGAACGCACTGGAGAGGGGTGCGGACCCAGTG 720
Qy 721 CTCTACTCGCGGAGAGTTTACCGAGTACGCTTGGCGCTGTACCAACAGTACCAAC 780
Db 721 CTCTACTCGCGGAGAGTTTACCGAGTACGCTTGGCGCTGTACCAACAGTACCAAC 780
Qy 781 CTGGCGGAGCACTACGCTTCGCGCAAGCTTATGGTGGCGTCTGCTTGGCTCTCTCC 840
Db 781 CTGGCGGAGCACTACGCTTCGCGCAAGCTTATGGTGGCGTCTGCTTGGCTCTCTCC 840
Qy 841 AACGTGCTCTCTCCAGCGCGCGCCCGCTACGAGGCGCTGCACTGTCACCAACCGGA 900
Db 841 AACGTGCTCTCTCCAGCGCGCGCCCGCTACGAGGCGCTGCACTGTCACCAACCGGA 900

Qy 901 GCCTTCGCGCTCTTGGGGTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 960
Db 901 GCCTTCGCGCTCTTGGGGTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 960
Qy 961 CTCGCGCTAGGCTCTTCCCGGCTTCCACTCAGTACGAGCGCGCGCTTCTGGGTCAAGCTG 1020
Db 961 CTCGCGCTAGGCTCTTCCCGGCTTCCACTCAGTACGAGCGCGCGCTTCTGGGTCAAGCTG 1020
Qy 1021 GCACCGGCGTCTGCTGCTTCTTCTCGGAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 GCACCGGCGTCTGCTGCTTCTTCTTCTCGGAGGGCGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Qy 1081 CCAGCGCTCTTCCGACCCCTTCTGGACCAAGCGCCCAAGGACTGCGAGCCAGGAGAGGG 1140
Db 1081 CCAGCGCTCTTCCGACCCCTTCTGGACCAAGCGCCCAAGGACTGCGAGCCAGGAGAGGG 1140
Qy 1141 GGTCTACCTCTTATCTTCCGCGACCCACTGCAAGCAGCGCGCTTCTCCAGACTTAAAA 1200
Db 1141 GGTCTACCTCTTATCTTCCGCGACCCACTGCAAGCAGCGCGCTTCTCCAGACTTAAAA 1200
Qy 1201 TGTATCACCACTTAACCTGTGAGGGGACCCCAATCTTGGACTTCTTCCCGCTTGGGACAT 1260
Db 1201 TGTATCACCACTTAACCTGTGAGGGGACCCCAATCTTGGACTTCTTCCCGCTTGGGACAT 1260
Qy 1261 CGCAGCGCGGGAAGTGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Db 1261 CGCAGCGCGGGAAGTGTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Qy 1321 AGCGCTGCTGCGCGAGCGCTTCCGACATCCGAGGCAACCAAGGAAAGTCTCTTGGGCGA 1380
Db 1321 AGCGCTGCTGCGCGAGCGCTTCCGACATCCGAGGCAACCAAGGAAAGTCTCTTGGGCGA 1380
Qy 1381 TCTGTAATAAAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1440
Db 1381 TCTGTAATAAAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1440
Qy 1441 A 1441
Db 1441 A 1441

RESULT 2
HSM806241 1740 bp mRNA linear PRI 17-JUN-2003
LOCUS Homo sapiens mRNA; CDNA DKFZp686C04213 (from clone DKFZp686C04213).
DEFINITION BX537581
ACCESSION BX537581.1 GI:31873659
VERSION
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1740)
AUTHORS Bloembergen, H., Boecher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobbe, G., Han, M. and Wiemann, S.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr. 1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBP (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686C04213) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/
location/Qualifiers
1. 1740
/organism="Homo sapiens"
/mol_type="mRNA"

```

/db_xref="taxon:9606"
/map="15q15"
/clone="DKFp686C04213"
/tissue_type="human rectum tumor"
/clone_1ib="586 (synonym: hlcc3). Vector psport1_sai; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
1694. .1699
polvA signal
polvA_site
1720
ORIGIN

```

| | | | | |
|-----------------------|--------------|--------------------|----------|-------------|
| Query Match | 71.4% | Score 1029 | DB 9 | Length 1740 |
| Best Local Similarity | 98.8% | Pred. No. 4.5e-207 | | |
| Matches 1038 | Conservative | Mismatches 15 | Indels 0 | Gaps 0 |

381 CTTGCCCGGGATCCGCTGCCCACTCGCGCTGGTTTTGGTTGATGAGAGTTCTTCTCAGTCT 440

688 CTGACCCGGGTGCATTCCTGCAAGCGCTGGTTTGGTTGGTAGAGTCTCTTCAGTCT 747

[illegible]

494
GTTCTAGGGCGAGAAAATGGTCCCTGTGCCTTCACCTCAGATTCTCACT

748 GTTCATAGGCGCAGAAATGTTGGCTGTGTCACCTCAGTGCAGAAATGTTCTGTCGGGACAGGAT 500

501 GAAACCAACACATCTTCAAAGCCTTCAGCGAGCGCGGTACAGCCGGTGCTCT 560

Db 808 GAACACCAACACATCTTACAAAGCCTTACGCCAGCGGGCGTTACAGCCCGTGCGGTCT 867

561 GCCTGGGCGCTGGAGGSCATTAATATTACACTCAGGGACCCGAGTGCATCAGTGAA 620

Db 868 GCTCGTGGCCCTGGAGGGCATTAAATTACACTCAGGGACCCAGTGCATCAGCTGAA 927

621 CGAGACCATTTGACTACAAACGAGCAGTTCACCTGGGCTCTGAAGAGAATTACGCCCGGA 680

928 CGAGGCCATTGACTACACGAGCAGTTCACTGGCGTCTGAAGAGAAATTACGCCGCGGA 987

52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
88

[illegible]

DB 988 GTACGGACGCAC TGGAGAGGGGCTG CCGGACCCAGATGCCTACCTGGCGAGTGGTCTT

741 CACACCGAGTAGCCCTTGGGCGCTGTACCAACAGTACACCTGGGGGGACACATACGCCGC 800

D_b 1048 CACACCGAGTAGCCCTTGGCGCCTGTACCAACAGTACCACTGGCGGACACTACGCCCTC 1107

QY 801 GGCCACGCTATGGGTGGCGTCTCTGCTTCTGGCTCTCTCCAAAGTGGTGCTCTCCACGCC 860

Db
1108 GGCCACGCTATGGGTGGCGTTCTGCTTCGGCTCTCTCCAACGTGTGCTCTCCAGGC 1167

861 GGCCCCGGTCTACGGAGGCCCTGGCACTGCTGACACCGGAGCCTTCGGCTCTTCGGGTT 920

db 1168 GGCCCGCTCTACGGAGGCTTGGCACTGCTACACCGGAGCCCTTCGGGCTCTTCGGGT 1227

Q: 921 CTTTCGGCTTGGCTTCCATTCCTCTAGGGTCGGCTCCGGCTAGGCTCCTCCGC 980

[illegible][illegible]

981 GCTCACCACCTCAGTACGGCGCGGCTTCCTAGGTCAAGTGGCAACCGGCGTCCCTGGTCCCT
QY

D5 1288 GCTCACCACCTCAGTACGGGGCGGCTTCTGGGTCAAGCTGGCAACCGGGGTCCTGTGCCCT 1347

Qy 1041 CTTCCCTCGAGGGCCGTCGTTGAGTCTCCAGTATGTTGGGCCCCAGGGCTCTTCGACCCCT 1100

Dh 1348 CTTCTCGAGGGCGTGGTGAATCCAGTATGTTGGCCACAGCGCTTTCACCT 1407

QY 1101 TCTGGACCAAGGCCAAGGACTGAGCCAGGAGAGAGGGGGCTCACCTCTTATCCTCGG 1160

db 1408 TCTGGACCAAGCGCCCAAGGACTGCAGCCAGGAGAGGGGGCTCACCTCTATCTCTCGG 1467

1161 CCACCCACTGTCACAGCAGCGCGGTCCCGAGACCTAAATGTATCAACCTAACCTGTG 1220

[illegible][illegible]

1221 AGGGGGACCCATCTGGACTCCCTCCCGGCTTGGGCAATCGACGCGCGGAGGAGAGC

Series: IPAK Plate: 58 Row: g Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13385245.

```

FEATURES             source
Location/Qualifiers
1. .1354
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db xref="taxon:10090"
/c1one="MGC:35715 IMAGE:4989815"
/tissue_type="Colon, normal, 5 month old male mouse."
/c1one_l1b="NCI CGAP_C024"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
1. .1354
/gene="9030623N16R1k"
/db xref="LocusID:66811"
/db xref="MGI:1914061"
129. .1091
/codon_start=1
/product="RIKEN cDNA 9030623N16"
/protein_id="AAH3111.1"
/db xref="GI:21411408"
/db xref="LocusID:66811"
/t_ranslation="MTAWDGLVFPYQPQPHRAASPSVPLTLVILVPLSLAASFLLPG
IRGRSWMFVLRVLLSLFVIAEYVAHFGSDVFGVRWNTSYKAFSPSRVQVHGVLH
LAGLVNLTURGTIPQOLNLTIDYNERFTWLNEDYTFYVALEKGLPDPDLVLAKE
PTPSSPCGLHYOHLVHAAATLVAFQFWIIANALSMAPALYGLAALTTCGATL
EGVPRFASISVPLCHPLRGASVLTYPYGAOSFWLTATGLISLLGLGGAWVILHYTRFS
KATGRIKQVQENQAKNCRGSLTYPNOHQIKSDINITTL"

```

| ORIGIN | Query Match | 49.4% | Score 711.4 | DB 10 | Length 1354 |
|--------|-----------------------|--|--------------------|-----------|-------------|
| | Best Local Similarity | 75.9% <td>Pred. No. 6.6e-140</td> <td></td> <td></td> | Pred. No. 6.6e-140 | | |
| | Matches 936 | Conservative 0 | Mismatches 281 | Indels 16 | Gaps 4 |
| QY | 222 | CTGCCGCGCTCGTGCAGCACTCGCGCGCGGTGCAGCATCACCCCTGTGGAAACGGCGTACT | 281 | | |
| DB | 92 | CTTCGCTGCTCGACAGTTCGGAATCTTTCGAGATGATCGCTGGGAGACGGGGTGTCT | 151 | | |
| QY | 282 | GGCTTTTACCCAGCCCGCGATGCCGAGCGCTTCAGAGTTTCCACTGCTCATGTTAT | 341 | | |
| DB | 152 | ACCCTTTTACCCCGACCGCAGCATGCCGCGAGCTTCAGGCTACCGTCTCATTTGTTAT | 211 | | |
| QY | 342 | TCTAGTGTTTTGGTCTTAGCAGCAAGCTTCTGCTCATTTTCGCGGGGAGATCGTGGCCA | 401 | | |
| DB | 212 | CCGTGGTTCTTGTTCTTTGGCTGTAGCTTCTCTTTCATCTTCCTCGGAATCCGTGGGCA | 271 | | |
| QY | 402 | CTCGCGCTGTTTGGTGGTTCAGAGTCTTCTTCAGTCTGTTCATAGGCGGCAGAAATGCT | 461 | | |
| DB | 272 | CTCGCGCTGTTTGGTGGTTCAGAGTCTTCTTCAGTCTGTTCATAGGCTGTTTCACTAGTTCGCGAATGCT | 331 | | |
| QY | 462 | GGCTGTGCACTTCAGTGCAGAAATGTTGTCGGGTACAGTGAACACCAACATCTCAACA | 521 | | |
| DB | 332 | GGCTGTGCACTTCAGTGCAGAAATGTTGTCGGGAGAGTGTGGACCAACATCTCAACA | 391 | | |
| QY | 522 | AGCCTTCAGCGCAGCGCGGTTTACAGCCGCTGTGGTCTGCTCGTGGGCGCTCGAGGCGCAT | 581 | | |
| DB | 392 | AGCCTTCAGTCCATCCGTGTTTCAAGTCCATGTGGTCTGCACTGTCGGGCTGGCGGCGT | 451 | | |
| QY | 582 | TAAATATTACACTACAGGGAACCCAGTCGATCAGTGAACGAGACCATTTGACTACAAAGA | 641 | | |
| DB | 452 | TAAATATTACACTCCGAGGAACACCCGAGGACAGCTGAAACGAGACCATTTGACTACAAAGA | 511 | | |
| QY | 642 | GCAGTTTCACTCGTGGTCTCAAGAGAAATTCGCGCGGAGTACGCGAACCCACTTCGAGAA | 701 | | |
| DB | 512 | GCCTTTTCACTTGGGCTCTGAACGAGACTACACCAAGGAGTATGTCATGCTTCGAGAA | 571 | | |
| QY | 702 | GGGCGTCCCGGACCCAGTGTCTTACCTCGCGGAGAGTGTTCACCGAGTAGTCCCTTGGCG | 761 | | |
| DB | 572 | GGGACTGCCGGAACCCAGTGTCTTACCTGCGAGAGAGTGTTCACCGAGTAGTCCCTTGGCG | 631 | | |
| QY | 762 | CTCTGATCCACAGTACACACTCGGCGGGAACACTTACGCTTCGCGCACGCTATGCGGTGGCGTT | 821 | | |

[illegible]

| | |
|------------|--|
| RESULT | 4 |
| AF547266 | |
| LOCUS | Sus scrofa dual oxidase 1 {DuoX1} mRNA, complete cds. |
| DEFINITION | AES47266 linear MAM 03-APR-2003 |
| ACCESSION | AF547266 |
| VERSION | AF547266.1 GI:236644372 |
| KEYWORDS | . |
| SOURCE | Sus scrofa {pig} |
| ORGANISM | Sus scrofa |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; |
| AUTHORS | Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus. |
| TITLE | 1 (bases 1 to 5851) Morand,S., Chaaraoui,M., Kaniewski,J., Deme,D., Ohayon,R., Noel-Hudson,M.S., Virion,A. and Dupuy,C. Effect of iodide on nicotinamide adenine dinucleotide phosphate oxidase activity and DuoX2 protein expression in isolated porcine thyroid follicles |
| JOURNAL | Endocrinology 144 (4), 1241-1248 (2003) |
| MEDLINE | 22526370 |
| PUBMED | 12639306 |
| REFERENCE | 2 (bases 1 to 5851) Kaniewski,J., Morand,S., Noel-Hudson,M.-S., Ohayon,R., Virion,A. and Dupuy,C. |
| AUTHORS | Direct Submission |
| TITLE | Submitted (13-SEP-2002) Unite 486, INSERM University of Paris-Sud, |
| JOURNAL | |

5 rue JB Clement, Chatenay-Malabry 92296, France
Location/Qualifiers
1. .5851
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
1. .5851
/gene="Duox1"
948. .5609
/gene="Duox1"
/EC number="1.6.3"
/functions="putatively generates H2O2 by transferring
electrons from NADH/NADPH to molecular oxygen"
/function="oxidoreductase acting on NAD(P)H with oxygen as
acceptor"
/note="NADPH:O2 oxidoreductase"
/codon_start=1
/product="dual oxidase 1"
/protein_id="AA39338.1"
/db_xref="GI:2366433"
/translation="MGFRLALAWLLVGPMPGARSISWEVDFDGVNMLMEHKW
GSKGRQLRVPAADGADGVYPLGEPHLPNDRDLSNMRGPAQASLRNRTVLGVPF
GYHLSDLVSTLEKCPAEFLNIHIPGDPVDFPHKSGDVLVLPQSRWDNPTQSPF
NPRDITNWTGWLDAISGSHSWSDBLRSFGQSLASGDPAPPAQADPLFMWTP
PDPAQTQWGPQGLVAGAEQGNRPFLQALGLLWFLYHNLCAQRLAHPQADGDEELF
QHAKRVATQSIITMYWLPFLKMPQEVTVYRPPDPSISPEFLAASQFFSTWV
PFGYMRNASCHPQVINEINSVSALVCNYSRSHENLQRAEDVALLGLMASOI
AEREDHVVVDVQDFWGPPLAFSRDHLASCLQGRDGLRSLPYTKARALGLPVTRW
QDINPALSRDGIIVLEAALYNQLSLELLPGLLESYSGPGLFSTIVLDQFVRL
RDGRYFENTNGFLSKETABITRSLRDLVAVNTMTPGALQPNVFFHAGDPCP
KRLQVONQSIKCEKLVGMKALEQWQKKEPCRPVLHLQSGQIHMDGRSLVRLTIQ
LRPPOVNLISLNSHGRITLLKI PKBYDLVLMFDEBERQVMNLSQALSKESGLSF
QEWELRQELMRAAVTQRSHLETFFHFLFSQVLDIQQADAGALPDSQKVRAL
TCELRAEFASLGLKQPMQFVSMFSLADKNGCYLSFRFELDLIVVMKSPREKS
RLMPFYDFDGNLSDLEFTQLCVKGAVVEIKDCRRASYSIOELKCPSPFVSAHC
LTMEDFPHMLRDLSELTQVQVGRVVEIKDCRRASYSIOELKCPSPFVSAHC
PSNVDFVEVLPFWKLGCTDTPDPEIRREPKVTSFQPLLFTEAREHFKPGRSH
QTVQOKRPFVNRHIGCLAVFYTIAGLFLERYAYFAAHMGITDITTVAGVILS
RGTASISPMFSVILLTMCRNLI TFLRFLNRYVYVFDAAVDFHLLIASTALITLVH
SAGHVNYLFSISPLSCLFGLFHDNGSEPPQKYVMFQVPLGLTGMVLLIL
AIMYFASHFRPCFSGVWVTHLILYLLIYLLIYLLIYLLIYLLIYLLIYLLIYLLIY
VGDVLSLRSHKVELSVKAEVLPVSTHQLQRFQGFYKSGVYKSGVYKSGVYKSGVY
HPEFTLSAPHEVLSLRHAGPWTTRLEIYSPPTDNCAPKLYLDGPGEGHQE
WHKFSVLSVGGVITPPASILKDLVFKSSVSCQVCKKIYFIWVTRQKQFELAD
IIRVEENDHRLDVSHIYITQAEKFLDRLTMYICRHFQKVLNRLSLFTGLRSLTH
EGEPPEPFNSLQEVHPQVRKIGVFSQGPQGMKXKACQLINRQDRTHPSHYEN
P"

FEATURES
source
gene
CDS

ORIGIN
Query Match 43.5%; Score 626.2; DB 4; Length 5851;
Best Local Similarity 81.8%; Pred. No. 7.6e-122;
Matches 735; Conservative 0; Mismatches 163; Indels 1; Gaps 1;
254 GCAGATGACCTGTGGAAGGGTACTGCTTTTACCCAGCCCGGCGATGCGCAG 313
10 GCAGATGACTCTGTGGAAGGGTCTGCTTTTACCCAGCTGCTGATGCGCAG 69
314 GCTTCAGCGTTCACCTGCTCATCTGTTATTCAGTGTGTTTGGCTTACGACGATCC 373
70 GCTTCAGCGTTCACCTGCTCATCTGTTATTCAGTGTGTTTGGCTTACGACGATCC 129
374 TGCTCATCTGCGGGATCCGTCGACCTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTT 433
130 TACTCATCTTGGCCGGATTCGTCGACCTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTT 189
434 TCAGTCTGTTTCATAGCGGAGAAATGTGGCTGTGCACTTCAGTCAGAAATGTTCTGG 493
190 TCAGCTGTTTCATAGCGGAGAAATGTGGCTGTGCACTTCAGTCAGAAATGTTCTGG 249
494 GTACAGTGAACACCAACATCTTACAAAGCTTTCAGCGAGCGCGGTTACAGCCGCTG 553
250 GCGAAGTTAGCAGCAATACATCTTACAAAGCTTTCAGTCAGTGTGTCAGCGACCC 309

QY 554 TCGTCTGCTCGTGGCGCTGAGGGCATTATATTAACCTACAGGACCCAGTGCATC 613
DB 310 TCGTCTGCAACGCTGGGGCTTGGAGGGTGTGAATATACACTACAGGAAATCCGCTACGC 369
QY 614 AGCTGAACAGAGACCAATTGACTACAAAGAGCAGTTTCACTTGGGCTGTCAAGAGAAATACG 673
DB 370 AGCTGAATGAGACCATCGACTACAAAGAGCAGTTTCACTTGGGCTGTATGGGGAATATG 429
QY 674 CCGGGAGTACGGAACGCACTGGAAGAGGGCTGGGACCCAGTGTCTTCTACCTGGCGG 733
DB 430 CCGGGAGTACGGAAGGGCTTCTGGAAGAGGGCTTGGGACCCAGTGTCTTCTACCTGGCGG 489
QY 734 AGAAGTTTCAACCCAGTAGCCCTTTCGCGGCTGTGACCAACAGTACCACTTGGCGGAGCACT 793
DB 490 AGAAGTTTACTCCAGAGAGCCCTTTCGCGGCTTATACCGGCTATATCGCTAGCCGAGCACT 548
QY 794 AGCTCGGCAACGCTATGAGTGTGCTTCTTCTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCT 853
DB 549 AGCTCGGCAACATATGAGAGGGCTTCTGGAAGAGGGCTTGGGACCCAGTGTCTTCTTCTTCT 608
QY 914 TCGGGGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 973
DB 569 TCTCTGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 728
QY 974 CTTCCGGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 1033
DB 729 CTTCCGGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 788
QY 1034 TGTGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 1093
DB 789 TGTGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 848
QY 1094 GCACCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 1152
DB 849 GCACCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 907
RESULT 5
AX035348 498 bp DNA linear PAT 15-NOV-2000
LOCUS Sequence 3 from Patent WO0053748.
DEFINITION AX035348
ACCESSION AX035348
VERSION AX035348.1 GI:11191065
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.
TITLE Novel compounds
JOURNAL Patent: WO 0053748-A 3 14-SEP-2000;
BRUCK CLAUDE ELVIRE MARIE (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ;
VINALS Y DE BASSOLS CARLOTTA (BE) ; COCHE THIERRY (BE) ; CASSART
JEAN POL (BE)
FEATURES
location/Qualifiers
1. .498
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
source
ORIGIN
Query Match 32.3%; Score 465.4; DB 6; Length 498;
Best Local Similarity 99.4%; Pred. No. 6.6e-88;
Matches 488; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 939 CTCTAGCTGCGCTCTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 997

[illegible]

| | | | |
|------------|--------|--|--------|
| Df | 157896 | GCGCTCTCCGGGTGTCAGGCAGCCCCCAGCTTGCTGGCTTGCCTGCCCGCTGCGTGAAC | 157955 |
| Qy | 241 | ACTCGGCGCGGTGTCAGCATGACCTCTGTGAAACGGGTGCTGCTTTTATCCCGAGCCC | 300 |
| Df | 157956 | ACTCGGCGCGGTGTCAGCATGACCTCTGTGAAACGGGTGCTGCTTTTATCCCGAGCCC | 158015 |
| Qy | 301 | CAGCATGCCGAGCTTCCAGGTTCCACTGCTCATGCTATTCTAGTGTGTTTTGGCTCTA | 360 |
| Df | 158016 | CAGCATGCCGAGCTTCCAGGTTCCACTGCTCATGCTATTCTAGTGTGTTTTGGCTCTA | 158075 |
| Qy | 361 | GCAGCAAGCTTCTGCTCATCTTTCGCGGGAGTCCGTGGCCACTGCGCGCTGTTTT | 415 |
| Df | 158076 | GCAGCAAGCTTCTGCTCATCTTTCGCGGGAGTCCGTGGCCACTGCGCGCTGTTTT | 158130 |
| RESULT 7 | | | |
| AC091117 | | 181312 bp DNA linear PRI 20-FEB-2002 | |
| LOCUS | | Homo sapiens chromosome 15 clone RP11-109D20 map 15q15, complete | |
| DEFINITION | | sequence. | |
| ACCESSION | | AC091117.5 GI:18767395 | |
| VERSION | | HTG. | |
| KEYWORDS | | Homo sapiens (human) | |
| SOURCE | | Homo sapiens | |
| ORGANISM | | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | |
| REFERENCE | | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| AUTHORS | | 1 (bases 1 to 181312) | |
| TITLE | | Rowen,J., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S., | |
| JOURNAL | | Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R., | |
| REFERENCE | | Pate,D. and Hood,L. | |
| AUTHORS | | Sequencing of human chromosome 15 DIS146-D15S117 region | |
| TITLE | | Unpublished | |
| JOURNAL | | 2 (bases 1 to 181312) | |
| REFERENCE | | Rowen,J., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S., | |
| AUTHORS | | Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R., | |
| TITLE | | Pate,D. and Hood,L. | |
| JOURNAL | | Direct Submission | |
| REFERENCE | | Submitted (30-MAR-2001) Multimegabase Sequencing Center, Institute | |
| AUTHORS | | for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA | |
| TITLE | | 98105, USA | |
| JOURNAL | | 3 (bases 1 to 181312) | |
| REFERENCE | | Rowen,J., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S., | |
| AUTHORS | | Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R., | |
| TITLE | | Pate,D. and Hood,L. | |
| JOURNAL | | Direct Submission | |
| REFERENCE | | Submitted (13-FEB-2002) Multimegabase Sequencing Center, Institute | |
| AUTHORS | | for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA | |
| TITLE | | 98105, USA | |
| JOURNAL | | 4 (bases 1 to 181312) | |
| REFERENCE | | Rowen,J., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S., | |
| AUTHORS | | Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R., | |
| TITLE | | Pate,D. and Hood,L. | |
| JOURNAL | | Direct Submission | |
| REFERENCE | | Submitted (14-FEB-2002) Multimegabase Sequencing Center, Institute | |
| AUTHORS | | for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA | |
| TITLE | | 98105, USA | |
| JOURNAL | | 5 (bases 1 to 181312) | |
| REFERENCE | | Rowen,J., Madan,A., Qin,S., Baradarani,L., Birditt,B., Bloom,S., | |
| AUTHORS | | Burke,J., Dors,M., Fleetwood,P., Kaur,A., Madan,A., Nesbitt,R., | |
| TITLE | | Pate,D. and Hood,L. | |
| JOURNAL | | Direct Submission | |
| REFERENCE | | Submitted (20-FEB-2002) Multimegabase Sequencing Center, Institute | |
| AUTHORS | | for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA | |
| TITLE | | 98105, USA | |
| JOURNAL | | On Feb 20, 2002 this sequence version replaced gi:18653547. | |
| COMMENT | | ----- Genome Center | |
| | | Center: Multimegabase Sequencing Center | |
| | | Center code: UWMS | |
| | | Web site: http://chroma.mbt.washington.edu/msg_www | |
| | | Contact: leorows@systemsbiology.org | |
| | | ----- Summary Statistics | |
| | | Sequencing vector: pUC18; L08752 | |

----- Genome Center -----
Center: Multimegabase Sequencing Center
Center code: UWMSC
Web site: http://chroma.mbt.washington.edu/mag_www
Contact: leorowen@systembiology.org
----- Summary Statistics -----
Sequencing vector: pUC18; L08752

Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-primer Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399

Note: Data from overlapping clone AC090888 [Drafting center: WIGR] was added for finishing. RP11-109D20 spans bases 1-174236 of this sequence. Bases 174237-179373 derive from clone RP11-276K9, AC087790 [Drafting center: WIGR] to establish the overlap with the next BAC in our tiling path. Bases 179373-181312 derive from finished clone CTD-2651B20, AC051619 [Drafting center: UWMSC] in order to give overlap required for determining long range contiguity.

| FEATURES | | Location/Qualifiers | |
|--------------|--|---------------------|--|
| source | 1. | 181312 | |
| | /organism="Homo sapiens" | | |
| | /mol_type="genomic DNA" | | |
| | /db_xref="taxon:9606" | | |
| | /chromosome="15" | | |
| misc_feature | 1. | 89137 | |
| | /note="overlap with CTD-2014N11, AC090888" | | |
| | 36107. | 36120 | |
| | /note="low quality data" | | |
| | 40470. | 40820 | |
| misc_feature | 53030. | 53480 | |
| | /note="sequence data generated from subcloned PCR product" | | |
| | 86982. | 86983 | |
| | /note="sequence data generated from subcloned PCR product" | | |
| | 95610. | 96030 | |
| misc_feature | 98330. | 97420 | |
| | /note="sequence data generated from subcloned PCR product" | | |
| | 98260. | 98420 | |
| | /note="sequence data generated from subcloned PCR product" | | |
| | 98850. | 98852 | |
| misc_feature | 100981. | 101003 | |
| | /note="low quality data" | | |
| | 174237. | 179372 | |
| | /note="low quality data" | | |
| | /note="overlap with RP11-276K9, AC087790. This sequence data was added to establish the overlap between the BACS in the tiling path" | | |
| misc_feature | 179373. | 181312 | |
| | /note="overlap with CTD-2651B20, AC051619. This data is added to give overlap for long range contiguity." | | |
| | | | |
| | | | |
| | | | |

ORIGIN

| | | | | |
|-------------|--|--------------|--------------------|----------------|
| Query Match | 28.1% | Score 405.4; | DB 9; | Length 181312; |
| | Best Local Similarity | 98.6%; | Pred. No. 4.8e-75; | |
| | Mismatches | 0; | Indels | 0; |
| | Conservative | 0; | Mismatches | 6; |
| | Caps | 0; | | |
| 1 | AAAGTAACGGCTACAGACAGTGAATAGTTTCGTCGCGCGCTAGAAAATCTGTG 60 | | | |
| 165016 | AAAGTAACGGCTACAGACAGTGAATAGTTTCGTCGCGCGCTAGAAAATCTGTG 165075 | | | |
| 61 | GTACCAACCCAGAGCGCTTGAGAGCAGCCACCTCCACGTTCTCTTAACGAGAGTGCA 120 | | | |
| 165076 | GTACCAACCCAGAGCGCTTGAGAGCAGCCACCTCCACGTTCTCTTAACGAGAGTGCA 165135 | | | |
| 121 | GGACTCAGATTACAGCCCACTGCTCCAGCGCTTGACGAAAGACGCGCAGGAC 180 | | | |
| 165136 | GGACTCAGATTACAGCCCACTGCTCCAGCGCTTGACGAAAGACGCGCAGGAC 165195 | | | |
| 181 | GGCTCTCCCGTCCAGGAGCCAGCTTGCTGGCTTGCCTGCCCGCTGGGTGAGC 240 | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

| | | |
|----|--------|--|
| Db | 165196 | GGCTCTCCCGCTCCAGGAGCCAGCTTGCTGGCTTGCCTGCCCGCTGGGTGAGC 165255 |
| Qy | 241 | ACTCGCGCGCGCTGACAGCATGACCTGTGTGAACGGGTACTGCTTTTACCCGAGCCC 300 |
| Db | 165256 | ACTCGCGCGCGCTGACAGCATGACCTGTGTGAACGGGTACTGCTTTTACCCGAGCCC 165315 |
| Qy | 301 | CGCATGCGCGAGCTTCACGGTTCCACTCTCATCTGTTATCTAGTGTGTTTGGCTCTA 360 |
| Db | 165316 | CGCATGCGCGAGCTTCACGGTTCCACTCTCATCTGTTATCTAGTGTGTTTGGCTCTA 165375 |
| Qy | 361 | GCAGCAAGCTTCCTGCTCATCTTTCGCGGGGATCGTGCGCCACTCGCGCTGTTT 415 |
| Db | 165376 | GCAGCAAGCTTCCTGCTCATCTTTCGCGGGGATCGTGCGCCACTCGTGAGGGTGT 165430 |

RESULT 8
AC09700/c

| | | | | | |
|------------|--|------------|-----|--------|-----------------|
| LOCUS | AC009700 | 156534 bp | DNA | linear | HTG 20-APR-2000 |
| DEFINITION | Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT | | | | |
| SEQUENCE | 15 unordered pieces. | | | | |
| ACCESSION | AC009700 | | | | |
| VERSION | AC009700.4 | GI:7622346 | | | |
| KEYWORDS | HTG; HTGS PHASE1; HTGS_DRAFT. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 156534) | | | | |
| JOURNAL | Birren, B., Linton, L., Nusbaum, C. and Lander, E. | | | | |
| REFERENCE | Homo sapiens chromosome 15, clone RP11-163P10 | | | | |
| AUTHORS | 2 (bases 1 to 156534) | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, J., Barna, N., Becker, R., Benn, J., Brown, A., Castle, J., Cerny, J., Collins, S., Collins, S., Collymore, A., Cooke, P., DeArilano, K., Depayre, E., Devon, K., Dewar, K., Donelan, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G., Haggis, B., Heaford, A., Horton, L., Howland, J. C., Jones, C., Kann, L., Karatas, A., Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J., Moll, M., Morris, W., Morrow, J., Mychaleckyj, J., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M. | | | | |

TITLE
JOURNAL

Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 20, 2000 this sequence version replaced gi:6056272.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
ftp://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu

----- Project Information
Center project name: L2336
Center clone name: 163 P10
----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 108837 bases at least Q40
Consensus quality: 130230 bases at least Q30
Consensus quality: 145227 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 155134; sum-of-contigs
Quality coverage: 3.2 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS
1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and Masuho, Y.
TITLE Full-length cDNAs
JOURNAL Patent: EP 1293569-A 1195 19-MAR-2003;
Helix Research Institute (JP); Research Association for Biotechnology (JP)
FEATURES
source
1. .2684
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 27.5%; Score 396; DB 6; Length 2684;
Best Local Similarity 97.6%; Pred. No. 3.5e-73;
Matches 402; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1007 TCTGGGTACGCTGGCAACGGGCTCTCTGCTCTTCTCTGAGGGGGCGGTGGAGTC 1065
Db 2363 TTTTCATCCCGCCACCGCCACAGGGGCTCTGCTCTTCTCTGAGGGGGCGGTGGAGTC 2304
Qy 1067 TCCAGTATGTTGGCGCCAGCGCTCTGCGACCTCTGCGACCAAGCGCCAGGACTGCA 1126
Db 2303 TCCAGTATGTTGGCGCCAGCGCTCTTGGACCTCTTGGACCAAGCGCCAGGACTGCA 2244
Qy 1127 GCCAGGAGAGGGGCTCAGCTCTTATCTCTGGGACCCACTGCAAGCAGCGCGCTC 1186
Db 2243 GCCAGGAGAGAGGGGCTCAGCTCTTATCTCTGGGACCCACTGCAAGCAGCGCGCTC 2184
Qy 1187 TCCAGACTTAAATCTGATACACACTTAACTCTGAGGGGGCCCACTCTGAGCTCTTCC 1246
Db 2183 TCCAGACTTAAATCTGATACACACTTAACTCTGAGGGGGCCCACTCTGAGCTCTTCC 2124
Qy 1247 CCGCTTGGGACATCGAGCGCGGGAGAGCAGTGCCTGCGAGCGCTGGGCGAGGAGCTC 1306
Db 2123 CCGCTTGGGACATCGAGCGCGGGAGAGCAGTGCCTGCGAGCGCTGGGCGAGGAGCTC 2064
Qy 1307 CAGGAGGCGCAGTACGCTGCTGGCGGAGCGCTGCGACATCCGAGGCGCCAGGAAA 1366
Db 2063 CAGGAGGCGCAGTACGCTGCTGGCGGAGCGCTGCGACATCCGAGGCGCCAGGAAA 2004
Qy 1367 GTCTCTGGGGGATCTGTAATAAACCTTTTCTTTTGTGTTTTTAAAAA 1418
Db 2003 GTCTCTGGGGGATCTGTAATAAACCTTTTCTTTTGTGTTTTTAAAAA 1952

RESULT 11
AK056896/c
LOCUS AK056896 Homo sapiens cDNA FLJ23234 fis, clone PROST2005426.
DEFINITION Homo sapiens cDNA FLJ23234 fis, clone PROST2005426.
ACCESSION AK056896
VERSION AK056896.1 GI:16552419
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS
1 Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S., Ono, Y., Hotta, T., Hiraoka, S., Murakawa, K., Takiguchi, S., Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M., Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi-Fujii, A.,

Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2684)
Isogai, T., Otsuki, T. and Sugiyama, T.
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES
source
1. .2684
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PROST2005426"
/tissue_type="prostate"
/clone_lib="PROST2"
/notes="cloning vector: pME18SFL3"
56. .952
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71304.1"
/db_xref="GI:16552420"
/translation="MTALGHTFFVYAGPKTPPMDTTLASIIIMIFLATALTIVILPGIKGKTLFLLRVVTSLSIFGAAISLFTGQVQNETINYNEEFTWRLGENTYAEYKALEKGLPDLVLAERKTPRSGLYRQVRLAGHTYSAMLVAFLLCWLLANVLSMPVLVYGGMLLATGTFOLLALFPMSATSLTSPCLHGLASVLHTHSGPAFWIPLTTLGLCLVLGLMAVAFRMPHRLKAFPNQSVDEDPLENSPEBGLSLPRYSMDSPSKSDIPLSEASSTKAYCKENHPKDPICAL"
ORIGIN
Query Match 27.5%; Score 396; DB 9; Length 2684;
Best Local Similarity 97.6%; Pred. No. 3.5e-73;
Matches 402; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 1007 TCTGGGTACGCTGGCAACGGGCTCTGCTCTTCTCTGAGGGGGCGGTGGAGTC 1066
Db 2363 TTTTCATCCCGCCACCGCCACAGGGGCTCTGCTCTTCTCTGAGGGGGCGGTGGAGTC 2304
Qy 1067 TCCAGTATGTTGGCGCCAGCGCTCTTGGACCCCTTCTGGACCAAGCGCCAGGACTGCA 1126
Db 2303 TCCAGTATGTTGGCGCCAGCGCTCTTGGACCCCTTCTGGACCAAGCGCCAGGACTGCA 2244
Qy 1127 GCCAGGAGAGGGGCTCAGCTCTTATCTCTGGGACCCACTGCAAGCAGCGCGCTC 1186
Db 2243 GCCAGGAGAGAGGGGCTCAGCTCTTATCTCTGGGACCCACTGCAAGCAGCGCGCTC 2184
Qy 1187 TCCAGACTTAAATCTGATACACACTTAACTCTGAGGGGGCCCACTCTGAGCTCTTCC 1246
Db 2183 TCCAGACTTAAATCTGATACACACTTAACTCTGAGGGGGCCCACTCTGAGCTCTTCC 2124
Qy 1247 CCGCTTGGGACATCGAGCGCGGGAGAGCAGTGCCTGCGAGCGCTGGGCGAGGAGCTC 1306
Db 2123 CCGCTTGGGACATCGAGCGCGGGAGAGCAGTGCCTGCGAGCGCTGGGCGAGGAGCTC 2064
Qy 1307 CAGGAGGCGCAGTACGCTGCTGGCGGAGCGCTGCGACATCCGAGGCGCCAGGAAA 1366
Db 2063 CAGGAGGCGCAGTACGCTGCTGGCGGAGCGCTGCGACATCCGAGGCGCCAGGAAA 2004
Qy 1367 GTCTCTGGGGGATCTGTAATAAACCTTTTCTTTTGTGTTTTTAAAAA 1418
Db 2003 GTCTCTGGGGGATCTGTAATAAACCTTTTCTTTTGTGTTTTTAAAAA 1952


```

QY 1067 TCAGATATGTTCCGCCAGCGCTCTTCGACACCCCTTCGACACCAAGAGGACTGCA 1126
Db 96171 TCAGATATGTTCCGCCAGCGCTCTTCGACACCCCTTCGACACCAAGAGGACTGCA 96230
QY 1127 GCCAGAGAGAGGGGCTCAGCTCTATCTCTGGGACCCACCTGACACAGCGCGCTC 1186
Db 96231 GCCAGAGAGAGGGGCTCAGCTCTATCTCTGGGACCCACCTGACACAGCGCGCTC 96290
QY 1187 TCCAGACTTAAATGTATCACCACCTAACCTGTGAGGGGACCCCAATCTGCACTCTTCC 1246
Db 96291 TCCAGACTTAAATGTATCACCACCTAACCTGTGAGGGGACCCCAATCTGCACTCTTCC 96350
QY 1247 CGCCTTGGGACATCGACGCGGGAGCAGTGCCTGGGACCCAGCGCTGGGCGAGAGCTC 1306
Db 96351 CGCCTTGGGACATCGACGCGGGAGCAGTGCCTGGGACCCAGCGCTGGGCGAGAGCTC 96410
QY 1307 CAGGAAGGCGCTGAGCGCTGTGCGCGGAGCGCTCGGACATCGCAGCGACCGAGGAAA 1366
Db 96411 CAGGAAGGCGCTGAGCGCTGTGCGCGGAGCGCTCGGACATCGCAGCGACCGAGGAAA 96470
QY 1367 GTCTCTGGGCGATCTGTAATAAACCTTTTCTTTTGTGTTTAAAA 1418
Db 96471 GTCTCTGGGCGATCTGTAATAAACCTTTTCTTTTGTGTTTAAAA 96522

RESULT 13
AX405919 1636 bp DNA linear PAT 14-JUN-2002
LOCUS Sequence 334 from Patent WO222660.
DEFINITION AX405919
ACCESSION AX405919.1 GI:21439277
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Tang, Y.T., Liu, C., Zhou, P., Asundi, V., Zhang, J., Zhao, Q.A., Ren, F.,
Xue, A.J., Yang, Y., Wehrman, T. and Drmanac, R.T.
TITLE Novel nucleic acids and polypeptides
JOURNAL Patent: WO 022660-A 334 21-MAR-2002;
HYSEQ, INC. (US)
FEATURES
source
1..1636
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
441..1472
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD35026.1"
/db_xref="GI:21439278"
/db_xref="EMBL:CAD35026"
/translation="MATLGTFFPYAGPKPTFMDTTLASIIIMFLATATFVILPG
IRKTRLPMLRVVTSLSFGAIALVAFNSVSVQSTNTSYKPSSEMSIDIGLQ
VGLGVNITLTPTVQOLNETINNEBFTWRLGENYAEYAKLEKGLPPVLYLTK
FTPRSPGLSYRLAGHYVTGAMLVAFNLWLANVMSLPALVGGVMLLATGIPOL
LALLFSPGATSLTSPCLHLGASVLVTHHGFATWTLTGLLVLLGLAMAVAHMOP
HRLKAFNQSVDEDPWLENSPEHGLLSPRVMSADSPKSDIPLSEASSTKAYCKEA
HPKDPDCAL"

Query Match 24.7%; Score 356.4; DB 6; Length 1636;
Best Local Similarity 63.9%; Pred. No. 86-65;
Matches 556; Conservative 0; Mismatches 311; Indels 3; Gaps 1;

QY 255 CAGCATGCCCTGTGGACGGGTACTGCTTTTACCCCGCCCGGATGGCGAGG 314
Db 437 CAGATGGCTACTTGGGACACACATTCCTCTTATGCTGGCCCGCCAGCCAACTTCCC 496
QY 315 CTTCAGCGTCCACTGCTCATCGTTATTCTAGTGTGTTTGGCTCTAGCAGCAAGCTTCT 374

```

```

Db 497 GATGGACACCACTTTGGCCAGCATCATCATGATCTTTCTGACTGCACTGGCCACGTTTCA 556
QY 375 GCTCATCTTTCGCGGCGATCCGTGGCCACCTCGCGCTGGTTTGGTGGTGAGAGTTCTTCT 434
Db 557 CGTCATCTTTCGCGGCGATTCGGGGAAGAGAGGCTGTTTGGCTGCTTCGGGTGCTGAC 616
QY 435 CAGTCTGTTATAGGCGCAAAATTTGGTGTGCTGCACTTCAGTGCAGAAATGTTGTTGGG 494
Db 617 CAGCTTATTTCATCGCGGCTGCAATCTCGCTGTGAAATTTCACTTCAGTGTGTTGTTGG 676
QY 495 TACAGTGAACACCAACACATCTTACAAAGCTTTCAGCGCAGCGCGCTTACACCGCGTGT 554
Db 677 CAGGTTCAGGACCAACACATCATCAAGGCTTCAGTCTGAGTGGATCAGCGCTGATAT 736
QY 555 CGTCTGCTGTGGGCTCTGGAGGCAATTAATATTACATCAACAGGACCCCGAGTGCATCA 614
Db 737 TGGGCTGCGAGGTTCGGGCTGGGTGGAGTCAACATCACTCACAGGACCCCGCTGCGACA 796
QY 615 GCTTGAACGAGACCAATTCATCAACGAGCAGTTTCACCTGGCGCTCTGAAAGAGAATTA 674
Db 797 GCTGAATGAGACCTCAATTAACACGAGGAGTTTCACTTGGCGCTGGGTGAGACTATGC 856
QY 675 CGCGAGTACCGGACGCACTGGAGAGGGGCTTCGCGACCCAGTGTCTTACCTGGCGGA 734
Db 857 TGAGGAGTATGCAAGGCTCTGGAGAGGGGCTTCGAGAGCCCTGTGTTGTACCTAACTGA 916
QY 735 GAAGTTTCAACCGAGTAGCCCTTTCGGCGCTGTACCAACAGTACCACTGCGCGGACACTA 794
Db 917 GAGTTTCACTCCAGAGGCCATGTGGCTTATACCGCAGTACCGCTTGGCGGACACTA 976
QY 795 CGCTCGGCGACGCTATGGTGGCGTTCGTCTGTGGCTTCCTTCAAGCTGCTGCTCTC 854
Db 977 CACCTCAGCCATCTATGGGTGGCATTCCTCTGCTGGCTGGCCAAATGTATGCTCTC 1036
QY 855 CAGCGCGCGCGCTCTTACGAGGCTTGGCACTGTGACCAACCGGAGCCTTCGCGCTCTT 914
Db 1037 CATGCTTGGCTGTATATGTTGGCTATGTTGGCTATGTTGGCTATGTTGGCTATGTT 1096
QY 915 CG---GGGTCTGCTTGGCGCTTCCATCTCTAGCGTGGCGCTTCTGCGCGCTCGCGTAGG 971
Db 1097 GGCTCTGCTCTTCTTCTTCCATGCGCACATCACTCACCTCACCTGTGCGCTGCGCTGAG 1156
QY 972 CTCCTCGCGCTCACCACTCAGTAGCGCGCGCTTCTGGGTCAGCTGGGCAACCGCGCT 1031
Db 1157 CGTCTGTGCTGCTACTTACCATGGGCTGCTCTTGTGATCATTTGACACAGGACT 1216
QY 1032 CTTGTGCTCTTCTTCTCGAGGGCGGTGGTGTAGTCTTCCAGTATGTTGGCGCCAGCGCTCT 1091
Db 1217 GCTGTGTGTGCTGCTGGGCTTGGCTATGCGCGTGGCGCCACAGGATGACAGCTCACAGGCT 1276
QY 1092 TCGCACCTTCTGGACCAAGCGCCAAAGGA 1121
Db 1277 GAAGGCTTCTTCAACACGAGTGTGATGA 1306

RESULT 14
BC029819 1923 bp mRNA linear PRI 06-OCT-2003
LOCUS Homo sapiens hypothetical protein FLJ32334, mRNA (cDNA clone
DEFINITION MGC:35310 IMAGE:5177619), complete cds.
ACCESSION BC029819
VERSION BC029819.1 GI:20987583
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1923)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

```

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, J.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnier, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2338257
JOURNAL MEDLINE
PUBMED
2477332
2 (bases 1 to 1923)
Strausberg, R.
Direct Submission
Submitted (06-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgpbbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Namavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 51 Row: 0 Column: 3
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21389332.

FEATURES

source

1..1923
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:35310 IMAGE:5177619"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..1923
/gene="FLJ32334"
/db_xref="LocusID:90527"
401..1852
/codon_start=1
/product="FLJ32334 protein"
/protein_id="AAH29819.1"
/db_xref="GI:20987584"
/db_xref="LocusID:90527"

gene

CDS

/translation="MATTGHTPPVAGPKPTPMDTTLASIIIMIFLTALATPVLPLG
IRGTFELFELAVTSLFTGAALLAVNPSSWSVGSTNTSYKAPSSSEWISADIGLQ
VSLGNNITLTGTPVQQLNETINYNEEFTWLGENYAEYAKALEKGLDPVLYIAEK
FTPRSPGQYRQLAGHTYSAMLWAFICWLLANVMSMPVLYTGGVYKLLATGIFOL
LALLPFSMATSLTSPCLPHLGSVLTHHGPAFVLTITGLCCVLGLAMAVAHMQP
HLKAPFNQSDVDPMLEWSPEGGLLSPYRSMADSPKSDIPILSEASSTKAYRPR
RLSLVPAVRLPALPGLALLAQAWRALLPGLRCPKAGKESRLGPPHSPMPFGP
EGCEERWAEHTGDSRPLSRGTGRLWRWGSKERRACGVRAMLPRLVNSGLKRPSCIL
DLPKCDYRDRAPFHLLEPTPCVTSRHTPLI"

| | | |
|------------|-----------------------|---|
| ORIGIN | Query Match | 24.7%; Score 356.4; DB 9; Length 1923; |
| | Best Local Similarity | 63.9%; Pred. No. 8.1e-65; |
| | Matches | 556; Conservative |
| QY | 255 | CAGCATGACCTGTGGAACGGCGTACTGCTTTTACCCCGGCGCATGCCGAGG 314 |
| DB | 397 | CAAGATGGGTACTTTGGGACACACATTCCTTCATGTGGCCCAAGCCHACCTTCCC 456 |
| QY | 315 | CTTCAGCGTTTCCACTGCTCATCGTTTATTCTAGTGTGTTTTTGGCTCTAGCAGCAAGCTTCCT 374 |
| DB | 457 | GATGGACACCATTTTGGCCGACCATCATCATGATCTTTCTGACTGCACTGCCACCGTTTCAT 516 |
| QY | 375 | GCTCATCTTGGCGGGGATCCGTGGCCACTCGCGCTGGTTTGGTTGGTGGTGGTGGTGGTGGTGGT 434 |
| DB | 517 | CGTCATCTGCTGCGCATTTGGGGAAAGACGAGGCTGTTTGGCTGCTTCGGGTGGTGGTGGTGGT 576 |
| QY | 435 | CAGTCTGTTTCATAGGCGCAGAAATTTGCTGCTGTGCACTTTCAGTGCAGAAATGGTTCGGGG 494 |
| DB | 577 | CAGCTTATTTCATCGGGGCTGCTATCTGCTGCTGTGAATTTTCAGTCTGAGTGGTCTGTGGG 636 |
| QY | 495 | TACAGTGAAACACCAAC 554 |
| DB | 637 | CCAGGTTCAGCACCAAC 696 |
| QY | 555 | CGGTCTGCTGCTGGGCTCGAGGGGCAATTAATATTACACTCACAGGACACCCAGTGCATCA 614 |
| DB | 697 | TGGGCTGCGAGTCCGGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 756 |
| QY | 615 | GCTGAACGAGACCAATTTGACTTACAAAGCCTTCAGCGGACGCGGCTTACAGCCCGTGT 674 |
| DB | 757 | GCTGAATGAGACCAATCAATTTACAAAGGAGTTTCACTTGGCGCTGGGTGGTGGTGGTGGTGGT 816 |
| QY | 675 | CGCGGAGTACCGACACGCACTGGAGAGGGCTGCGGACCCAGTGCCTTACCTGGCGGA 734 |
| DB | 817 | TGAGGAGTATGCAAAAGGCTCTGGAGAGAGGGCTGCGGACCCAGTGCCTTACCTGGCGGA 876 |
| QY | 735 | GAAGTTTCAACACCAAGTACGCTTTCGGGCTGTGTACCAACAGTACCACTGCGCGGAGCACTA 794 |
| DB | 877 | GAAGTTTCACTCCAGAGAGCCCATGTGSCCTATACCGCCAGTACCGCTGGCGGAGCACTA 936 |
| QY | 795 | CGCTCGGCGCACCGTATGGTGGGTGCTGCTTCTGCTGCTGCTTCTCCAAAGTGGTGGTGGTGG 854 |
| DB | 937 | CACCTACGACCATGCTATGGGTGGCATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996 |
| QY | 855 | CACCGCGCGCGCTCTACGAGGCGCTGGCACTGCTGACCAACCGGAGGCTTTCGGGCTCTT 914 |
| DB | 997 | CATGCTGTGCTGGTATATGTTGGCTACATGCTATTTGSCCAACGCGGATCTTCCAGCTGT 1056 |
| QY | 915 | CG---GGGTCTTTCGCTTGGCTTCCATCTCTAGGGTGGCTGCTGCGCGCTGCGCGCTAGG 971 |
| DB | 1057 | GGCTCTGCTCTTCTTCTTCTTCCATGGCCACATCACTCACCTCAACCTTCTCCCTGCACTGGG 1116 |
| QY | 972 | CTCCTCGCGCTCACCACTAGTACGGCGCGCTTCTGGGTACGCTGCGCAACCGCGCGT 1031 |
| DB | 1117 | CGCTTCTGCTGTGATTAATCACTGATGGCTGCTGCTTCTGATCACATTGACACAGGACT 1176 |
| QY | 1032 | CCTGTGCTCTTCTTCGAGGGGCGGTGGTGGAGTCTTCAGTATGTTTGGCGCCAGGCTCT 1091 |
| DB | 1177 | GCTGTGTGCTGCTGCTGGCGCTGCTATGGCGGTGGCGCCACAGGATGCAGGCTCAGGCT 1236 |
| QY | 1092 | TCGCACCTCTTCGACCAACAGCGCCAAAGGA 1121 |
| DB | 1237 | GAAGGCTTCTTCAACACAGAGTGGATGA 1266 |
| RESULT 15 | | |
| BC019755 | | |
| LOCUS | | |
| DEFINITION | BC019755 | 1521 bp mRNA linear ROD 04-OCT-2003 |
| ACCESSION | BC019755 | Mus musculus cDNA sequence BC019755, mRNA (cDNA clone MSC:25550 IMAGE:3964798), complete cds. |

| | | |
|-----------|---|-------------|
| VERSION | BC019755.1 | GI:18043594 |
| KEYWORDS | MGC. | |
| SOURCE | Mus musculus (house mouse) | |
| ORGANISM | Mus musculus | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1521) | |
| AUTHORS | Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.K., Rubin,G.M., Hong,L., Stappleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Frange,C., Raha,S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McSwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S.K., Garcia,A.M., Gay,I.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kertanen,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,N., Madan,A., Young,A.C., Shvachenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schermer,A., Schein,J.E., Jones,S.J. and Marra,M.A. | |
| TITLE | Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences | |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) | |
| MEDLINE | 22388257 | |
| PUBMED | 12477932 | |
| REFERENCE | 2 (bases 1 to 1521) | |
| AUTHORS | Strausberg,R. | |
| TITLE | Direct Submission | |
| JOURNAL | Submitted (19-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA | |
| REMARK | NIH-MGC Project URL: http://mgc.nci.nih.gov | |
| COMMENT | Contact: MGC help desk Email: cqabp@mail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: ang@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A. | |

| | |
|---|-----|
| db_xref="MG1:2384961" | CDS |
| 285_1310 | |
| /codon_start=1 | |
| /product="cDNA sequence BC019755" | |
| /protein_id="AAH19755.1" | |
| /db_xref="GI:18043595" | |
| /db_xref="LocusID:213696" | |
| /translation="MAALGHTLPFYFTGKTFPPMDTFLAVIITIFLTVLPIILPGLRGKTLFWLRVVTSLFTGAVTLA VNPSSGVGHVNNATNTKAFSPGWSDVGLQICLGGVNIITFTGTPVQLNETNINFAFWELGRSVAEYKALEKLPDPVLVLAEXFTFSPGCLTNQRLAGHTASAMLMVAFICLLANVMPLVYVGGHMLATGSPQLLALPFFSMTSLISPCPLRGTAVLTHRHGPAFWITLATGLLILGLLVMVAHRMQP HRLKAPFNLSDDVLEWGSBEGLLSPHYRSIAESPETQDTFMSVASSETCFKFBKHPKESDCSL" | |
| ORIGIN | |
| Query Match 24.2%; Score 348.8; DB 10; Length 1521; | |
| Best Local Similarity 63.5%; Pred. No. 3.2e-63; | |
| Matches 549; Conservative 0; Mismatches 312; Indels 3; Gaps 1; | |
| QY 279 ACTGCTTTTATACCCAGCCCGCGCATGCGCAGCGCTTCAGCGTTCACACTGCTCATCTCT 338 | |
| DB 305 ACTGCCCTTCTACACTGGCACCAGCCAACTTTCCAATGGACACTACTCTGGCTGTCA 364 | |
| QY 339 TATTCTAGTGTGTTTGGCTCTAGCAGCAAGCTTCTGCTCTCATCTTCGCGGGATCCGTGG 398 | |
| DB 365 TATCACCATTCTTCTGACTGCTGCTGCTCACCCTTCACTTATATCTGCTGCGCATCCGTG 424 | |
| QY 399 CCACCTCGCGCTGTTTGGTGTGAGAGTCTTCTCTCAGTCTCTTTCATAGGCGCCGAAAT 458 | |
| DB 425 GAAGACGAGGCTGTTCTGGCTGCTTCCGGTGGTGCACAGCTTATTCATCGGGCTGTGAT 484 | |
| QY 459 TGTGCTGTGCATCTCAGTCAGCAAGATGTTGCTGGGTACAGTGAAACACCAACATCTCTTA 518 | |
| DB 485 CTTGGCTGTGAACCTTCAGTCTGAGTGGTCTGTGGGGCCATGTCAACGCCAACACACGTA 544 | |
| QY 519 CAAAGCCTTCAGCGCAGCGCGGCTTACAGCCCGTGTGCTGCTGCTGGGCTCGCAGGG 578 | |
| DB 545 CAAAGCCTTCAGTCCCAAGTGGGTGCTGTGGAGCTGGAGCTGCAGTCCGGCTCGGGCG 604 | |
| QY 579 CATTAATATTACATCTACAGGGAACCCAGATGTCATCAGCTGAACAGACCAATTCATCAA 638 | |
| DB 605 CGTCAACATCACTTTCACAGGCAACCCAGTGCAGCAGCTGAATGAACCAACATCAACAA 664 | |
| QY 639 CGAGCAGTTCACCTGGGTCTGAAAAGAAATTACCGCGGAGTACGCGAACCACTGGA 698 | |
| DB 665 TGAGGCATTCGCATGCGCGCTGGGCGCATGTCAGAGAGAGTATGCAAAAGGCCCTTGA 724 | |
| QY 699 GAAGGGGCTGCGGAGACCCAGTGTCTTACCTGGCGGAGAGTTTACACCCAGTAGGCGCTTG 758 | |
| DB 725 GAAGGGCCTTCAGACCCAGTGTCTTACCTGGCTGGAGAGTTCACCCCTCGAAGCCCATG 784 | |
| QY 759 CGGCTGTACCAACAGTACCACTTGGCGGAGACACTAGCGCTCGGCTGCGCTATGCGTGGC 818 | |
| DB 785 TGGCCTGTACAAACAGTACCGCTTGGCAGGACACTATGCGCTCTGCGCATGCTGTGGGTGG 844 | |
| QY 819 GTTCTGCTTCTGGGTCTCTTCCAACTGCTGTCTTCCACCGCGGCCCGCTCTACCGAGG 878 | |
| DB 845 ATTCTCTGCTGGCTGTGCGCAATGATGTTGTGCAATGCCAGTGTGTTTATGTTGTTG 904 | |
| QY 879 CTGGCACTGCTGACCAACCGGAGCCTTGGCGCTTTCGCGGTCTTTCGCCCTTGGCC---TC 935 | |
| DB 905 CCACATGCTGTGGCCACTGGTCTCTTCCAGCTGTTGGCGCTGTTCTTCTTCTTCTCCATGAC 964 | |
| QY 936 GATCTCTAGCTGCGGCTCTGCGCGCTCCCGCTAGGCTCTCTCGCGCTCACCACTCACTA 995 | |
| DB 965 CACGTCATCTATATCGCCCTGTGCCCTTGGGCACTGCTGTGTCGACACTCACCA 1024 | |
| QY 996 CGCGCGCGCGCTTCTGGGTCACTGGCAACCGGCGTCTCTGTGCTCTTCTCTCGAGCGGGC 1055 | |
| DB 1025 TGGGCGCGCTTCTGGATCACTGTAGCCACAGGACTGCTGTGATTTCTGCTTGGCTGTG 1084 | |
| QY 1056 CTGTGTGAGTCTTCCAGTATGTTTCGGCCCGCGCTCTTTCGCAACCTTCTGCGACCAAGCGC 1115 | |

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 30 Row: e Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21703789.
Location/Qualifiers
1. .1521
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:25550 IMAGE:3964798"
/tissue type="Mammary tumor. Metallothionien-1GF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone lib="NCI CGAP_Mam1"
/lab host="DH10E"
/note="Vector: pCMV-SPORT6"
1. .1521
/gene="BC019755"
/note="synonym: MGC25550"
/db_xref="LocusID:213696"

gene

5
 4
 4
 5
 4
 4
 5
 5
 5

DB 1085 CATGGCAGTGGCCACAGGATGCAGCTCAGACTGAAGGCTTCTCAAGCTGATTC 1144

QY 1116 CRAAGGACTGCAGCCAGGAGGAGG 1139

DB 1145 AGAGGACCCAGTCTGGAGTGGGG 1168

Search completed: June 3, 2004, 11:57:53
Job time : 5819 secs

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

OX nucleic - nucleic search, using sw model

Run on: June 3, 2004, 11:58:00 ; Search time 5808 Seconds (without alignments)
10753.669 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- 1: gb_ba:*
- 2: gb_brg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|--------------------|
| 1 | 1441 | 100.0 | 1441 | 6 | AX035346 Sequence |
| 2 | 976 | 67.7 | 1740 | 9 | BS37581 Homo sapi |
| 3 | 425 | 29.5 | 498 | 6 | AX035348 Sequence |
| 4 | 405 | 28.1 | 166937 | 2 | AC012255 Homo sapi |
| 5 | 405 | 28.1 | 181312 | 9 | AC091117 Homo sapi |
| 6 | 392 | 27.2 | 2684 | 6 | AX714511 Sequence |
| 7 | 392 | 27.2 | 2684 | 9 | AK056896 Homo sapi |
| 8 | 392 | 27.2 | 156534 | 2 | AC009700 Homo sapi |
| 9 | 392 | 27.2 | 171444 | 2 | AC087790 Homo sapi |
| 10 | 354 | 24.6 | 156534 | 2 | AC009700 Homo sapi |
| 11 | 217 | 15.1 | 217 | 9 | HS183H12F |
| 12 | 166 | 11.5 | 1923 | 9 | BC029819 Homo sapi |
| 13 | 136 | 9.4 | 218 | 9 | HS183A12R |
| 14 | 39 | 2.7 | 1725 | 9 | BC004161 Homo sapi |
| 15 | 39 | 2.7 | 1726 | 9 | BC003157 Homo sapi |
| 16 | 38 | 2.6 | 1260 | 9 | BC017580 Homo sapi |
| 17 | 38 | 2.6 | 2165 | 6 | BD275921 Uncouplin |
| 18 | 38 | 2.6 | 2384 | 6 | AX780522 Sequence |
| 19 | 38 | 2.6 | 2384 | 6 | AX780523 Sequence |
| 20 | 38 | 2.6 | 3993 | 9 | BC036370 Homo sapi |
| 21 | 38 | 2.6 | 5642 | 6 | AX345771 Sequence |
| 22 | 38 | 2.6 | 10480 | 6 | AX347103 Sequence |
| 23 | 38 | 2.6 | 174790 | 2 | AC134964 Canis fam |
| 24 | 38 | 2.6 | 180059 | 2 | AC023928 Homo sapi |
| 25 | 38 | 2.6 | 180763 | 2 | AC067826 Homo sapi |
| 26 | 38 | 2.6 | 185951 | 2 | AC024386 Homo sapi |
| 27 | 38 | 2.6 | 186051 | 2 | AC015683 Homo sapi |
| 28 | 37 | 2.6 | 1426 | 10 | AF175224 Rattus no |
| 29 | 37 | 2.6 | 1640 | 9 | BC050738 Homo sapi |
| 30 | 37 | 2.6 | 1678 | 9 | BS37581 Homo sapi |
| 31 | 37 | 2.6 | 1724 | 9 | BC003101 Homo sapi |
| 32 | 37 | 2.6 | 2645 | 3 | AY118446 Drosophil |
| 33 | 37 | 2.6 | 3734 | 9 | BS37581 Homo sapi |
| 34 | 37 | 2.6 | 3989 | 9 | BS37581 Homo sapi |
| 35 | 37 | 2.6 | 3990 | 9 | BS37581 Homo sapi |
| 36 | 37 | 2.6 | 135628 | 2 | AC067764 Homo sapi |
| 37 | 37 | 2.6 | 133166 | 2 | AC048357 Homo sapi |
| 38 | 37 | 2.6 | 190371 | 2 | AC116830 Homo sapi |
| 39 | 37 | 2.6 | 349980 | 6 | AX344557 Trypanoso |
| 40 | 36 | 2.5 | 1059 | 3 | AF036416 Mus muscu |
| 41 | 36 | 2.5 | 1407 | 10 | BC013450 Mus muscu |
| 42 | 36 | 2.5 | 1603 | 10 | BC024516 Mus muscu |
| 43 | 36 | 2.5 | 1713 | 9 | BC003624 Homo sapi |
| 44 | 36 | 2.5 | 2155 | 9 | BC015749 Homo sapi |
| 45 | 36 | 2.5 | 2255 | 6 | AR168373 Sequence |

ALIGNMENTS

RESULT 1
AX035346
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AX035346
Sequence 1 from Patent WO0053748.
AX035346.1 GI:1191064

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL

Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.
Novel compounds
Patent: WO 0053748-A 1 14-SEP-2000;

AX035346 1441 bp DNA linear PAT 15-NOV-2000

Pred. No. is the number of results predicted by chance to have a

BRUCK CLAUDELINE ELVIRE MARIE (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ;
VINALS Y DE BASSOLS CARLOTTA (BE) ; COCHE THIERRY (BE) ; CASSART
JEAN POL (BE)

FEATURES source
Location/Qualifiers
1..1441
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1441; DB 6; Length 1441;
Best Local Similarity 100.0%; Prid. No. 0;
Matches 1441; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTAAGCGGTACAGACAGTGAAGATAGTTTCCTCGCGCGGTAGAGAAAACCTCTGCG 60
DB 1 AAAGTAAGCGGTACAGACAGTGAAGATAGTTTCCTCGCGCGGTAGAGAAAACCTCTGCG 60
QY 61 GTACCAACCCAGAGCGTTGAGAGAGCCACCTCCAGCTTCCTTAACGAGAGGTGCA 120
DB 61 GTACCAACCCAGAGCGTTGAGAGAGCCACCTCCAGCTTCCTTAACGAGAGGTGCA 120
QY 121 GGAATCAGACTTACAGACCCCACTGGTCCAGCCTTGTACGAAAGAGAGCCCAAGGAC 180
DB 121 GGAATCAGACTTACAGACCCCACTGGTCCAGCCTTGTACGAAAGAGAGCCCAAGGAC 180
QY 181 GCGCTCTCCGCGTCCAGAGCCCACTGGTCCAGCCTTGTACGAAAGAGAGCCCAAGGAC 240
DB 181 GCGCTCTCCGCGTCCAGAGCCCACTGGTCCAGCCTTGTACGAAAGAGAGCCCAAGGAC 240
QY 241 ACTCGCGCGGTGAGAGATGACCTGTGGAAGCGGCTGCTTTTACCCCGAGGCC 300
DB 241 ACTCGCGCGGTGAGAGATGACCTGTGGAAGCGGCTGCTTTTACCCCGAGGCC 300
QY 301 CGGCATGCGCGAGGCTTCAGAGTCCACTGCTCATCGTTATCTAGTGTGTTGCTCTA 360
DB 301 CGGCATGCGCGAGGCTTCAGAGTCCACTGCTCATCGTTATCTAGTGTGTTGCTCTA 360
QY 361 GCAGAGGCTTCTGCTCATCTGCGGGGATCGTGGCCACTCGGCTGTGTTGGTGG 420
DB 361 GCAGAGGCTTCTGCTCATCTGCGGGGATCGTGGCCACTCGGCTGTGTTGGTGG 420
QY 421 GTGAGAGTCTTCTCAGTCTGTTCATAGCGCGAGAAATGTGCTGTGCTACGTGCA 480
DB 421 GTGAGAGTCTTCTCAGTCTGTTCATAGCGCGAGAAATGTGCTGTGCTACGTGCA 480
QY 481 GAATGTTCTGTTGAGTACAGTGAACCAACACATCTTACAAAGCCCTTACGAGCGGCG 540
DB 481 GAATGTTCTGTTGAGTACAGTGAACCAACACATCTTACAAAGCCCTTACGAGCGGCG 540
QY 541 GTTACAGCCGCTGTGCTGTGCGGCTGAGGCGCAATTAATTAACACTCACAGG 600
DB 541 GTTACAGCCGCTGTGCTGTGCGGCTGAGGCGCAATTAATTAACACTCACAGG 600
QY 601 ACCCAGTGCATCAGTGAACGAGACCAATGACTACAGAGAGTTCACCTGGGCTCTG 660
DB 601 ACCCAGTGCATCAGTGAACGAGACCAATGACTACAGAGAGTTCACCTGGGCTCTG 660
QY 661 AAAGAGATTAAGCGCGAGTACGAGACGACCTGAGAGAGGCTGCGGACCCAGT 720
DB 661 AAAGAGATTAAGCGCGAGTACGAGACGACCTGAGAGAGGCTGCGGACCCAGT 720
QY 721 CTCTACCTCGCGAGAGTTCACCGAGTACGCTTGTGCGGCTGTGACCAAGTACAC 780
DB 721 CTCTACCTCGCGAGAGTTCACCGAGTACGCTTGTGCGGCTGTGACCAAGTACAC 780
QY 781 CTGCGGAGACACTACGCTCGGCGACGCTATGAGTGGGCTGCTGCTGCTGCTCTCC 840
DB 781 CTGCGGAGACACTACGCTCGGCGACGCTATGAGTGGGCTGCTGCTGCTGCTCTCC 840
QY 841 AACGCTGCTCTCTCCACGCGCGGCCCTCTACGAGGCTGCGCACTGCTGACCAACCGA 900
DB 841 AACGCTGCTCTCTCCACGCGCGGCCCTCTACGAGGCTGCGCACTGCTGACCAACCGA 900

QY 901 GCCTTCGCGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 960
DB 901 GCCTTCGCGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 960
QY 961 CTCGCGCTAGGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1020
DB 961 CTCGCGCTAGGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1020
QY 1021 GCAACCGGCGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1080
DB 1021 GCAACCGGCGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1080
QY 1081 CCCAGCGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1140
DB 1081 CCCAGCGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1140
QY 1141 GGTCTCACTCTTATCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1200
DB 1141 GGTCTCACTCTTATCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1200
QY 1201 TGTATCACTCACTTAACTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1260
DB 1201 TGTATCACTCACTTAACTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1260
QY 1261 CGCAGGCGCGGAGAGTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1320
DB 1261 CGCAGGCGCGGAGAGTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1320
QY 1321 AGCGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1380
DB 1321 AGCGCTCTTCGCGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 1380
QY 1381 TCTGTAATAAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1440
DB 1381 TCTGTAATAAATCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 1440
QY 1441 A 1441
DB 1441 A 1441

RESULT 2
HSM806241 1740 bp mRNA linear PRI 17-JUN-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp686C04213 (from clone DKFZp686C04213).
DEFINITION
ACCESSION BX537581
VERSION BX537581.1 GI:31873659
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1740)
Bloeker, H., Boescher, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobbe, G., Han, M. and Wiemann, S.
DIRECT SUBMISSION
Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
COMMENT
Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by GPC (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686C04213) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.
FEATURES
source
1..1740
/organism="Homo sapiens"
/mol_type="mRNA"

| | | | | | | | | | |
|--|------|---|------|--|--|--|--|--|--|
| <div>/db_xref="taxon:9606" /map="15q15" /clone="DRP2p686C04213" /tissue_type="human rectum tumor" /clone_lib="686 (synonym: hlcc3). Vector pSport1_sfi; host DH10B; sites SfiIIA + SfiIB" /dev stages="adult" polyA signal 1694..1699 polyA_site 1720 ORIGIN</div> | | | | | | | | | |
| Query Match 67.7%; Score 976; DB 9; Length 1740; Best Local Similarity 99.9%; Pred. No. 0; Matches 1026; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | | | | | | | | | |
| QY | 405 | GCGCTGGTTTGGTGGGAGATCTCTCAGTCGTTCATAGCGCAGAAATTTGGC | 464 | | | | | | |
| DB | 712 | GCGCTGGTTTGGTGGGAGATCTCTCAGTCGTTCATAGCGCAGAAATTTGGC | 771 | | | | | | |
| QY | 465 | TGTGACATTCAGTCAGAAATGTTGGTGGTACAGTGAACACCAACACATCTCAAGC | 524 | | | | | | |
| DB | 772 | TGTGACATTCAGTCAGAAATGTTGGTGGTACAGTGAACACCAACACATCTCAAGC | 831 | | | | | | |
| QY | 525 | CTTACGACGCGCGCTTACGCCGTGTGGTCTCTGTCGTGGCCCTGGAGGGGATTA | 584 | | | | | | |
| DB | 832 | CTTACGACGCGCGCTTACGCCGTGTGGTCTCTGTCGTGGCCCTGGAGGGGATTA | 891 | | | | | | |
| QY | 585 | TATTACATTCACAGGACCCAGTCATCAGCTGAACAGACCAATTGACTACACGACA | 644 | | | | | | |
| DB | 892 | TATTACATTCACAGGACCCAGTCATCAGCTGAACAGACCAATTGACTACACGACA | 951 | | | | | | |
| QY | 645 | GTTCACCTGGGCTCTGAAGAGAAATAGCCGCGAGTACGGGAACGCACTGGAGAGG | 704 | | | | | | |
| DB | 952 | GTTCACCTGGGCTCTGAAGAGAAATAGCCGCGAGTACGGGAACGCACTGGAGAGG | 1011 | | | | | | |
| QY | 705 | GCTGCGGACCGAGCTCTACCTGCGGAGAGTTACACCGAGTACCTTTGGGCT | 764 | | | | | | |
| DB | 1012 | GCTGCGGACCGAGCTCTACCTGCGGAGAGTTACACCGAGTACCTTTGGGCT | 1071 | | | | | | |
| QY | 765 | GTACCACAGTACACCTGGCGGGACACTAGCCCTCGGCACGCTATGGGTGGGCTTG | 824 | | | | | | |
| DB | 1072 | GTACCACAGTACACCTGGCGGGACACTAGCCCTCGGCACGCTATGGGTGGGCTTG | 1131 | | | | | | |
| QY | 825 | CTTCTGGTCTCTCCAGTGTCTCTCCAGCGGGCGGCTCTACGAGGCGCTGCG | 884 | | | | | | |
| DB | 1132 | CTTCTGGTCTCTCCAGTGTCTCTCCAGCGGGCGGCTCTACGAGGCGCTGCG | 1191 | | | | | | |
| QY | 885 | ACTGCTGACACCGAGCTCTCGGCTCTTCGGGCTTCCTGGGCTTCCTATCTCTAG | 944 | | | | | | |
| DB | 1192 | ACTGCTGACACCGAGCTCTCGGCTCTTCGGGCTTCCTGGGCTTCCTATCTCTAG | 1251 | | | | | | |
| QY | 945 | CGTCCGCTCTCGGCTCTCGGCTCTAGGCTCTCCGCGCTCAGCACTCAGTACGGCGCG | 1004 | | | | | | |
| DB | 1252 | CGTCCGCTCTCGGCTCTCGGCTCTAGGCTCTCCGCGCTCAGCACTCAGTACGGCGCG | 1311 | | | | | | |
| QY | 1005 | CTTCTGGTTCAGCTGGCAACCGGCTCTGCTCTTCCTCGAGGGCGGCTGGTGG | 1064 | | | | | | |
| DB | 1312 | CTTCTGGTTCAGCTGGCAACCGGCTCTGCTCTTCCTCGAGGGCGGCTGGTGG | 1371 | | | | | | |
| QY | 1065 | TCTCCAGTATGTTGGGCCAGGCTCTTCGCAACCTTCCTGGACCAAGCGCAAGCATG | 1124 | | | | | | |
| DB | 1372 | TCTCCAGTATGTTGGGCCAGGCTCTTCGCAACCTTCCTGGACCAAGCGCAAGCATG | 1431 | | | | | | |
| QY | 1125 | CAGCAGGAGAGGGGGCTCACCTCTTATCTCTCGGCAACCACTGCAACGACGCGCG | 1184 | | | | | | |
| DB | 1432 | CAGCAGGAGAGGGGGCTCACCTCTTATCTCTCGGCAACCACTGCAACGACGCGCG | 1491 | | | | | | |
| QY | 1185 | TCTCCAGACTTAAATGTATCACCACTAACCTGTGAGGGGACCAATCTGGACTCTT | 1244 | | | | | | |
| DB | 1492 | TCTCCAGACTTAAATGTATCACCACTAACCTGTGAGGGGACCAATCTGGACTCTT | 1551 | | | | | | |
| QY | 1245 | CCCGGCTTGGACATCGCAGCGCGGAGAGTGCCTGCGGCGAGGCTGGCCAGAGGAC | 1304 | | | | | | |

1552 CCOCGCTTGGGACATCGCAGCGCGGGAAGCAGTGCCTCGCCAGGCTTGGCCAGGAGAC 1611
1305 TCCAGGAGGGGCTAGCGCTCTCGCGCGAGGCTCTCGGACATCGCAGGACCCAGGGA 1364
1612 TCCAGGAGGGGCTAGCGCTCTCGCGCGAGGCTCTCGGACATCGCAGGACCCAGGGA 1671
1365 AAGTCTCTCGGCGGATCTGTAATAAACCTTTTCTTTTCTTTTAAAAA 1424
1672 AAGTCTCTCGGCGGATCTGTAATAAACCTTTTCTTTTCTTTTAAAAA 1731
1425 AAAAAA 1431
1732 AAAAAA 1738
RESULT 3
AX035348
LOCUS AX035348 498 bp DNA linear PAT 15-NOV-2000
DEFINITION Sequence 3 from Patent WO0053748.
ACCESSION AX035348
VERSION AX035348.1 GI:11191065
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
REFERENCE
AUTHORS Bruck,C.E., Vinals,Y.D., Coche,T. and Cassart,J.P.
TITLE Novel compounds
JOURNAL Patent: WO 0053748-A 3 14-SEP-2000;
BRUCK CLAUDE ELVIRE MARIE (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ;
VINALS Y DE BASSOLS CARLOTA (BE) ; COCHE THIERRY (BE) ; CASSART
JEAN POL (BE)
FEATURES
source 1..498
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 29.5%; Score 425; DB 6; Length 498;
Best Local Similarity 100.0%; Pred. No. 5; Se-224;
Matches 425; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 998 GCGCGCTTCTGGGTCAAGTGGCAACCGGCTCTCTCTCGGAGGGGCG 1057
DB 62 GCGCGCTTCTGGGTCAAGTGGCAACCGGCTCTCTCTCGGAGGGGCG 121
QY 1058 TGGTGAATCTCAGTATGTTGGCCCGAGCGCTCTTCGACCCCTTTGGACCAAGCGCA 1117
DB 122 TGGTGAATCTCAGTATGTTGGCCCGAGCGCTCTTCGACCCCTTTGGACCAAGCGCA 181
QY 1118 AGGACTCGACGAGGAGGGGGCTCAGCTCTTATCTCGGCGAGCCCTGCAACAGC 1177
DB 182 AGGACTCGACGAGGAGGGGGCTCAGCTCTTATCTCGGCGAGCCCTGCAACAGC 241
QY 1178 AGGCGCTCTCCAGACTTAAAAATGATATCACTAACTGTGGGGGAGCCCAATCTGG 1237
DB 242 AGGCGCTCTCCAGACTTAAAAATGATATCACTAACTGTGGGGGAGCCCAATCTGG 301
QY 1238 ACTCTTCCCGCTTGGGACATCGAGGCGGAGAGTGGCCGCGAGGCTTGGGCA 1297
DB 302 ACTCTTCCCGCTTGGGACATCGAGGCGGAGAGTGGCCGCGAGGCTTGGGCA 361
QY 1298 GGAGAGTCTCCAGGAGGGGCTAGCGCTGTGGCGGAGGCTCGGACATCCGCGAGCA 1357
DB 362 GGAGAGTCTCCAGGAGGGGCTAGCGCTGTGGCGGAGGCTCGGACATCCGCGAGCA 421
QY 1358 CCAGGAGAGTCTCCTGGGCGATCTGTAAATAAATCTTTTCTTTTGTAAAAA 1417
DB 422 CCAGGAGAGTCTCCTGGGCGATCTGTAAATAAATCTTTTCTTTTGTAAAAA 481
1418 AAAAA 1422

| TITLE | REFERENCE | AUTHORS |
|---|--|---------|
| Sequencing of human chromosome 15 | D15S146-D15S117 region | |
| Unpublished | | |
| 2 (bases 1 to 181312) | | |
| Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S., Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R., Pate, D. and Hood, L. | | |
| Direct Submission | | |
| Submitted (30-MAR-2001) | Multimegabase Sequencing Center, Institute for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA 98105, USA | |
| 3 (bases 1 to 181312) | | |
| Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S., Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R., Pate, D. and Hood, L. | | |
| Direct Submission | | |
| Submitted (13-FEB-2002) | Multimegabase Sequencing Center, Institute for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA 98105, USA | |
| 4 (bases 1 to 181312) | | |
| Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S., Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R., Pate, D. and Hood, L. | | |
| Direct Submission | | |
| Submitted (14-FEB-2002) | Multimegabase Sequencing Center, Institute for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA 98105, USA | |
| 5 (bases 1 to 181312) | | |
| Rowen, L., Madan, A., Qin, S., Baradarani, L., Birditt, B., Bloom, S., Burke, J., Dors, M., Fleetwood, P., Kaur, A., Madan, A., Nesbitt, R., Pate, D. and Hood, L. | | |
| Direct Submission | | |
| Submitted (20-FEB-2002) | Multimegabase Sequencing Center, Institute for Systems Biology, 4225 Roosevelt Way NE, Suite 200, Seattle, WA 98105, USA | |
| On Feb 20, 2002 this sequence version replaced gi:18653547. | | |
| ----- | | |
| Genome Center Sequencing Center | | |
| Center code: UWMSC | | |
| Web site: http://chroma.mbt.washington.edu/msg_wmw | | |
| Contact: leerowen@systemsbiology.org | | |
| ----- | | |
| Summary Statistics | | |
| Sequencing vector: pUC18: L08752 | | |
| Chemistry: Dye-terminator Big Dye; 90% of reads | | |
| Chemistry: Dye-primer Big Dye; 10% of reads | | |
| Assembly program: Phrap; version 0.990399 | | |
| Note: Data from overlapping clone AC090888 [Drafting center: WIBR] was added for finishing. RP11-109D20 spans bases 1-174236 of this sequence. Bases 174237-179373 derive from clone RP11-276K9. AC087790 [Drafting center: WIBR] to establish the overlap with the next BAC in our tiling path. Bases 179373-181312 derive from finished clone CTD-2651820, AC051619 [Drafting center: UWMSC] in order to give overlap required for determining long range contiguity. | | |
| Location/Qualifiers | | |
| 1. .181312 | | |
| /organism="Homo sapiens" | | |
| /mol_type="genomic DNA" | | |
| /db_xref="taxon:9606" | | |
| /chromosome="15" | | |
| /map="15c15" | | |
| /clone="RP11-109D20" | | |
| /clone_lib="RPC1 human BAC library 11" | | |
| /note="This clone overlaps CTD-2014N11 AC090888 and RP11-276K9 AC087790. Data from overlapping BACs were added and the consensus sequence determined from RP11-109D20 to the extent possible." | | |
| 1. .89137 | | |
| /note="overlap with CTD-2014N11, AC090888" | | |
| 36107. .36120 | | |
| /note="low quality data" | | |
| 40470. .40820 | | |
| misc_feature | | |
| misc_feature | | |

COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source Location/Qualifiers
1. 2684
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PRO2005426"
/tissue types="prostate"
/clone_lib="PROS2"
/notes="cloning vector: pME18SFL3"
56. .952
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71304.1"
/db_xref="GI:16552420"
/translation="MATLGHFFPYAGKPTFPMDDTLASIMFLTALATFIVILPG
IRKQTRFLWLRVTSFICAAILGTPVQQLNFTINNEFTWELGNVABEYAKALE
KGLPDVLYLAERFTSPQGLYRQYRLAGHYTSAMLVAFPLCMLLNVLSPLVIV
GGYMLLATGIFOLLALIFSMATSLATSPCLHLGASVLTHGPAFWITLTGLLCVL
LGLAWVAHMRMOPHRLKARFNQSDVEDPMLWSPBEGLLSPYRSWADSPKQDIPL
SEASSTKAYCKEAPKDPDCAL"

CDS
27.2%; Score 392; DB 9; Length 2684;
Query Match
Best Local Similarity 100.0%; Pred. No. 1.2e-205;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
Qy 1027 GGGTCCTGTCCTCTTCCTCGGAGGGCCGTGTGTAGTCTCCAGTATGTCGCCCCAGC 1086
Db 2343 GGGTCCTGTCCTCTTCCTCGGAGGGCCGTGTGTAGTCTCCAGTATGTCGCCCCAGC 2284
Qy 1087 GCTCTTGGACCCCTCTTGACCAAGCGCCAGGACTCGACCCAGGAGAGGGGGCTCA 1146
Db 2283 GCTCTTGGACCCCTCTTGACCAAGCGCCAGGACTCGACCCAGGAGAGGGGGCTCA 2224
Qy 1147 CCTTTATCTCGGAGCCCACTGCACAGCAGGCGGCTCTCCAGACTTAAATATGATC 1206
Db 2223 CCTTTATCTCGGAGCCCACTGCACAGCAGGCGGCTCTCCAGACTTAAATATGATC 2164
Qy 1207 ACCACTAACCTGTGAGGGGACCCCAATCTGGACTCTCTTCCCGCTTGGGACATCGCAGG 1266
Db 2163 ACCACTAACCTGTGAGGGGACCCCAATCTGGACTCTCTTCCCGCTTGGGACATCGCAGG 2104
Qy 1267 CCGGGAAGCAGTGCCTCGGAGGCTTGGCCAGGAGAGCTTCAGGAAGGCACTGAGCGCT 1326
Db 2103 CCGGGAAGCAGTGCCTCGGAGGCTTGGCCAGGAGAGCTTCAGGAAGGCACTGAGCGCT 2044
Qy 1327 GCTGGCGGAGGCTTCGACATCGCAGGACCCAGGGAAGTCTCCCTGGGCGCATCTGTA 1386
Db 2043 GCTGGCGGAGGCTTCGACATCGCAGGACCCAGGGAAGTCTCCCTGGGCGCATCTGTA 1984
Qy 1387 AATAAACCTTTTCTTTCTTTTGTATTTTAAAAA 1418
Db 1983 AATAAACCTTTTCTTTCTTTTGTATTTTAAAAA 1952

RESULT 8
AC009700
LOCUS
DEFINITION Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
AC009700
AC009700.4 GI:7622346
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 156534)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Birket,B., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Depayre,B., Devon,K., Dewar,K., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L., Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meidrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A., Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
2 (bases 1 to 156534)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Birket,B., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A., Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Depayre,B., Devon,K., Dewar,K., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L., Karatas,A., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meidrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B., Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P., Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A., Testaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 20, 2000 this sequence version replaced gi:6056272.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2336
Center clone name: 163 P.10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 108837 bases at least Q40
Consensus quality: 130230 bases at least Q30
Consensus quality: 145227 bases at least Q20
Insert size: 157000; agarose-1p
Insert size: 155134; sum-of-ctnigs
Quality coverage: 3.2 in Q20 bases; agarose-1p
Quality coverage: 3.3 in Q20 bases; sum-of-ctnigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1169: contig of 1169 bp in length
* 1170: gap of 100 bp
* 1270: contig of 2086 bp in length
* 3335: gap of 100 bp
* 3435: gap of 100 bp
* 5698: contig of 2263 bp in length
* 5798: gap of 100 bp
* 5799: contig of 3418 bp in length
* 9217: gap of 100 bp
* 9317: contig of 6317 bp in length
* 15633: gap of 100 bp
* 15733: contig of 6013 bp in length
* 21746: gap of 100 bp
* 21847: contig of 8505 bp in length
* 30351: gap of 100 bp
* 30352: contig of 6058 bp in length
* 36503: gap of 100 bp
* 36509: contig of 100 bp

```

* 36610 45279: contig of 8670 bp in length
* 45280 45379: gap of 100 bp
* 45800 56851: contig of 11472 bp in length
* 56852 56951: gap of 100 bp
* 56952 67374: contig of 10423 bp in length
* 67375 67474: gap of 100 bp
* 67475 79534: contig of 12060 bp in length
* 79535 95432: gap of 100 bp
* 95433 95432: contig of 15798 bp in length
* 95433 95433: gap of 100 bp
* 95533 120801: contig of 25269 bp in length
* 120802 120901: gap of 100 bp
* 120902 156534: contig of 35833 bp in length.

```

FEATURES

```

source
1..156534
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone_lib="RP11-11 Human Male BAC"
/clone="RP11-163P10"
1..1169
/notes="assembly_fragment"
1270..3335
/notes="assembly_fragment"
3436..5698
/notes="assembly_fragment"
5799..9216
/notes="assembly_fragment"
9317..15633
/notes="assembly_fragment"
15734..21746
/notes="assembly_fragment"
21847..30351
/notes="assembly_fragment"
30452..36509
/notes="assembly_fragment"
clone_end:T7
vector_side:left"
36610..45279
/notes="assembly_fragment"
45380..56851
/notes="assembly_fragment"
56952..67374
/notes="assembly_fragment"
67475..79534
/notes="assembly_fragment"
79635..95432
/notes="assembly_fragment"
95533..120801
/notes="assembly_fragment"
120902..156534
/notes="assembly_fragment"

```

ORIGIN

```

Query Match 27.2%; Score 392; DB 2; Length 156534;
Best Local Similarity 100.0%; Pred. NO. 1.2e-205;
Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1027 GCGTCTCTGCGCTCTCTCCGAGGGCGCGTGTGAGTCTCCAGTATGTTCGGCCCGCAGC 1086
Db 96131 GGGTCTCTGCGCTCTCTCCGAGGGCGCGTGTGAGTCTCCAGTATGTTCGGCCCGCAGC 96190
Qy 1087 GCTTTTCGACCCCTTTCTGGAACCAAGCGCCAGGACTGCGAGGAGAGGGGGCTCA 1146
Db 96191 GCTTTTCGACCCCTTTCTGGAACCAAGCGCCAGGACTGCGAGGAGAGGGGGCTCA 96250
Qy 1147 CCTTTATCTCTGGCGACCCCTGCGACAGCGCGCTCTCCAGACTTAAATCTATC 1206
Db 96251 CCTTTATCTCTGGCGACCCCTGCGACAGCGCGCTCTCCAGACTTAAATCTATC 96310
Qy 1207 ACCACTAACCTGTGAGGGGACCCCAATCTGGACTCTTCCCGCTTGGGACATCGCAGG 1266

```

```

Db 96311 ACCACTAACCTGTGAGGGGACCCCAATCTGGACTCTTCCCGCTTGGGACATCGCAGG 96370
Qy 1267 CCGGGAAGCAGTCCCGCCAGGCTTGGCCAGGAGAGCTCCAGGAAGGGCACTGAGCGCT 1326
Db 96371 CCGGGAAGCAGTCCCGCCAGGCTTGGCCAGGAGAGCTCCAGGAAGGGCACTGAGCGCT 96430
Qy 1327 GCTGGCGGAGGCTCTCGGACATCCGAGGCGACGAGGAAGTCTCTCGGGCGGATCTGTGA 1386
Db 96431 GCTGGCGGAGGCTCTCGGACATCCGAGGCGACGAGGAAGTCTCTCGGGCGGATCTGTGA 96490
Qy 1387 AATAAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1418
Db 96491 AATAAACCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 96522

```

```

RESULT 9
AC087790 171444 bp DNA linear HTG 09-MAY-2001
LOCUS Homo sapiens chromosome 15 clone RP11-276K9 map 15, WORKING DRAFT
DEFINITION SEQUENCE, 13 unordered pieces.
AC087790
AC087790.2 GI:13357344
HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 171444)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone RP11-276K9
Unpublished
2 (bases 1 to 171444)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barra,N., Bastien,V., Boguslavskiy,L., Bouhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Menus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-JAN-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 16, 2001 this sequence version replaced gi:12498501.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 112451
Center clone name: 276 K 9
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731

```


Consensus quality: 165906 bases at least Q40
 Consensus quality: 168587 bases at least Q30
 Consensus quality: 169557 bases at least Q20
 Insert size: 180000; agarose-efp
 Insert size: 170244; sum-of-contigs
 Quality coverage: 5.7 in Q20 bases; agarose-efp
 Quality coverage: 6.0 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 13 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 6091: contig of 6091 bp in length
 * 6092 6191: gap of 100 bp
 * 6192 7986: contig of 1795 bp in length
 * 7987 8086: gap of 100 bp
 * 8087 43939: contig of 35853 bp in length
 * 43940 44039: gap of 100 bp
 * 44040 45298: contig of 1259 bp in length
 * 45299 45398: gap of 100 bp
 * 45399 47595: contig of 2197 bp in length
 * 47596 47695: gap of 100 bp
 * 47696 52385: contig of 4690 bp in length
 * 52386 52485: gap of 100 bp
 * 52486 56310: contig of 3825 bp in length
 * 56311 56410: gap of 100 bp
 * 56411 71313: contig of 14903 bp in length
 * 71314 71413: gap of 100 bp
 * 71414 90993: contig of 13590 bp in length
 * 90994 91093: gap of 100 bp
 * 91094 114559: contig of 20366 bp in length
 * 114560 115559: gap of 100 bp
 * 115560 137226: contig of 26167 bp in length
 * 137227 137826: gap of 100 bp
 * 137827 170053: contig of 32227 bp in length
 * 170054 170154: gap of 100 bp
 * 170154 171444: contig of 1291 bp in length.

FEATURES

Source

1. 171444
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15"
 /clone="RP11-276K9"
 /clone_lib="RPC1-11 Human Male BAC"

1. 6091

/note="assembly_fragment"

clone_end:876

vector_side:left"

6192..7986

/note="assembly_fragment"

8087..43939

/note="assembly_fragment"

44040..45298

/note="assembly_fragment"

45399..47595

/note="assembly_fragment"

47696..52385

/note="assembly_fragment"

52486..56310

/note="assembly_fragment"

56411..71313

/note="assembly_fragment"

71414..90993

/note="assembly_fragment"

91094..114559

/note="assembly_fragment"

111560..137226

/note="assembly_fragment"

111560..137226

/note="assembly_fragment"

111560..137226

/note="assembly_fragment"

111560..137226

/note="assembly_fragment"

misc_feature /note="assembly_fragment"
 137827..170053
 misc_feature /note="assembly_fragment"
 170154..171444
 misc_feature /note="assembly_fragment"
 clone_end:17
 vector_side:right"

ORIGIN

Query Match 27.2%; Score 392; DB 2; Length 171444;
 Best Local Similarity 100.0%; Pred. No. 1.2e-205;
 Matches 392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1027 GGCCTCTCTGCTCTCTCTCTCGAGGGGCGTGGTGGTCTCCAGTATGTTGGCCCGCAGC 1086
 |||||
 Db 143780 GGGCTCTCTGCTCTCTCTCTCGAGGGGCGTGGTGGTCTCCAGTATGTTGGCCCGCAGC 143839
 |||||

Qy 1087 GCTCTTCGCAACCTCTCTGAGACCAAGCGCAAGGCTGACGAGGAGAGAGGGGGCTCA 1146
 |||||
 Db 143840 GCTCTTCGCAACCTCTCTGAGACCAAGCGCAAGGCTGACGAGGAGAGGGGGCTCA 143899
 |||||

Qy 1147 CTTCTTATCTCTCGCGACCACTGCAAGCAGCGCTCTCCAGACTTTAAATGTATC 1206
 |||||
 Db 143900 CTTCTTATCTCTCGCGACCACTGCAAGCAGCGCTCTCCAGACTTTAAATGTATC 143959
 |||||

Qy 1207 ACCACTAACCTGTGAGGGGAGCCCAATCTGCACTCTCCCGCTTGGGACATCGCAGG 1266
 |||||
 Db 143960 ACCACTAACCTGTGAGGGGAGCCCAATCTGCACTCTCCCGCTTGGGACATCGCAGG 144019
 |||||

Qy 1267 CCGGGAAGCAGTCCCGCCAGGCTCGGCGAGGAGCTCCAGGAAGGCACTGAGCGCT 1326
 |||||
 Db 144020 CCGGGAAGCAGTCCCGCCAGGCTCGGCGAGGAGCTCCAGGAAGGCACTGAGCGCT 144079
 |||||

Qy 1327 GCTGGCGGAGGCTCGGACATCGCGGACCAAGGAAAGTCTCTGGGGCGATCTGTA 1386
 |||||
 Db 144080 GCTGGCGGAGGCTCGGACATCGCGGACCAAGGAAAGTCTCTGGGGCGATCTGTA 144139
 |||||

Qy 1387 AATAAACTTTTCTTTTGTGTTTTTAAAAA 1418
 |||||
 Db 144140 AATAAACTTTTCTTTTGTGTTTTTAAAAA 144171
 |||||

RESULT 10

AC009700/c

LOCUS

AC009700 156534 bp DNA linear HTG 20-APR-2000
 Homo sapiens chromosome 15 clone RP11-163P10 map 15, WORKING DRAFT
 SEQUENCE, 15 unordered pieces.

AC009700

AC009700.4

GI:7622346

HTG; HTGS_PHASE1; HTGS_DRAFT.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 156534)

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 15, clone RP11-163P10

Unpublished

2 (bases 1 to 156534)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,

Baker,J., Baldwin,J., Barna,N., Beckerly,R., Benn,J., Brown,A.,

Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,

Cooke,P., DeArelano,K., Depayre,E., Devon,K., Dewar,K.,

Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C.,

Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,

Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,

Karatas,A., Lebecky,J., Lieu,C., Locke,K., Macdonald,P.,

Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,

Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J.,

Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,

Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,

Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,

Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,

Unpublished

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
Direct Submission
Submitted (28-AUG-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 20, 2000 this sequence version replaced gi:6056272.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu

Project Information
Center project name: L2336
Center clone name: 163_P_10

Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 108837 bases at least Q40
Consensus quality: 130230 bases at least Q30
Consensus quality: 145227 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 155134; sum-of-contigs
Quality coverage: 3.2 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1169: contig of 1169 bp in length
* 1170 1269: gap of 100 bp
* 1270 3335: contig of 2066 bp in length
* 3336 3435: gap of 100 bp
* 3436 5698: contig of 2263 bp in length
* 5699 5798: gap of 100 bp
* 5799 9217: contig of 3418 bp in length
* 9217 15633: contig of 6317 bp in length
* 15634 21746: gap of 100 bp
* 21747 21846: gap of 100 bp
* 21847 30351: contig of 8505 bp in length
* 30352 36509: contig of 6058 bp in length
* 36510 36609: gap of 100 bp
* 36610 45279: contig of 8670 bp in length
* 45280 45379: gap of 100 bp
* 45380 56851: contig of 11472 bp in length
* 56852 56951: gap of 100 bp
* 56952 67374: contig of 10423 bp in length
* 67375 67474: gap of 100 bp
* 67475 79534: contig of 12060 bp in length
* 79535 79634: gap of 100 bp
* 79635 95432: contig of 15798 bp in length
* 95433 95532: gap of 100 bp
* 95533 120801: contig of 25269 bp in length
* 120802 120901: gap of 100 bp
* 120902 156534: contig of 35633 bp in length.
* Location/Qualifiers
* 1. .156534
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="15"
* /map="15"
* /clone="RP11-163P10"

/clone lib="RP11-11 Human Male BAC"
1. .1169
/note="assembly_fragment"
1270. .3335
/note="assembly_fragment"
3436. .5698
/note="assembly_fragment"
5799. .9216
/note="assembly_fragment"
9317. .15633
/note="assembly_fragment"
15734. .21746
/note="assembly_fragment"
21847. .30351
/note="assembly_fragment"
30452. .36509
/note="assembly_fragment"
clone end: T7
vector side: left
36610. .45279
/note="assembly_fragment"
45380. .56851
/note="assembly_fragment"
56952. .67374
/note="assembly_fragment"
67475. .79534
/note="assembly_fragment"
79635. .95432
/note="assembly_fragment"
95533. .120801
/note="assembly_fragment"
120902. .156534
/note="assembly_fragment"

ORIGIN
Query Match 24.6%; Score 354; DB 2; Length 156534;
Best Local Similarity 99.8%; Pred. No. 1.4e-184;
Matches 404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAAGTAACGGCTACAGACAGTGGAGAAATAGTTTCGCTCGCCGCTAGAAAACTCTCTCG 60
Db 7148 AAAGTAACGGCTACAGACAGTGGAGAAATAGTTTCGCTCGCCGCTAGAAAACTCTCTCG 7089
Qy 61 GTACCAACCCAGAGCGTTGAGAGAGCCAGCCAGCTTCCAGCGTTCTTAACGAGAGGTGCA 120
Db 7088 GTACCAACCCAGAGCGTTGAGAGAGCCAGCCAGCTTCCAGCGTTCTTAACGAGAGGTGCA 7029
Qy 121 GGACTCAGACTTCCAGAGCCCACTCGGTCCAGCCTTGTTACGCAAAAGAGAGCGAC 180
Db 7028 GGACTCAGACTTCCAGAGCCCACTCGGTCCAGCCTTGTTACGCAAAAGAGAGCGAC 6969
Qy 181 GCGCTCTCCGCGTCCAGAGCCCACTCGGTCCAGCCTTGTTACGCTCCGCTCGGTGAGC 240
Db 6968 GCGCTCTCCGCGTCCAGAGCCCACTCGGTCCAGCCTTGTTACGCTCCGCTCGGTGAGC 6909
Qy 241 ACTCGCGCGGTGAGAGCCCACTCGGTGAGAGCCCACTCGGTGAGAGCCCACTCGGTGAGAGCC 300
Db 6908 ACTCGCGCGGTGAGAGCCCACTCGGTGAGAGCCCACTCGGTGAGAGCCCACTCGGTGAGAGCC 6849
Qy 301 CGGCATCCCGAGCGTTTCAGCGTTTCAGCGTTTCAGCGTTTCAGCGTTTCAGCGTTTCAGCGTT 360
Db 6848 CGGCATCCCGAGCGTTTCAGCGTTTCAGCGTTTCAGCGTTTCAGCGTTTCAGCGTTTCAGCGTT 6789
Qy 361 GCAGCAAGCTTCCGCTCATCTTCGCGGATCGGTGCGCACTCG 405
Db 6788 GCAGCAAGCTTCCGCTCATCTTCGCGGATCGGTGCGCACTCG 6744

RESULT 11
HS183H12P/c HS183H12P 217 bp DNA linear PRI 18-OCT-1995
LOCUS H.sapiens Cpg island DNA genomic MseI fragment, clone 183h12,
DEFINITION forward read cp9183h12.ft1a.

TITLE
JOURNAL
COMMENT
FEATURES
source

ACCESSION 257552
 VERSION 257552.1 GI:1028783
 KEYWORDS Cpg island; genomic MseI fragment.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Cross, S.H., Charlton, J.A., Nan, X. and Bird, A.P.
 TITLE Purification of CpG islands using a methylated DNA binding column
 JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
 MEDLINE 94282070
 PUBMED 8012384

REFERENCE 2 (bases 1 to 217)
 AUTHORS Dodsworth, S.J., Huckle, E., Wilkinson, P. and Micklem, G.
 TITLE Direct Submission
 JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk

COMMENT
 Vector: pGEM-52f(-)
 Clones are available from the UK MRC Human Genome Mapping Project Resource Centre, Hinxton, Cambridgeshire CB10 1RQ. UK. See URL: <http://www.hgmp.mrc.ac.uk/> for details
 or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES
 source
 1..217
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="183h12"
 /sex="male"
 /tissue_type="blood"
 /clone_lib="CGI-1"
 /dev_stage="adult"

ORIGIN
 Query Match 15.1%; Score 217; DB 9; Length 217;
 Best Local Similarity 100.0%; Pred. No. 1.4e-108;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Gy 1198 AAATGTTATCACCTTGTAGGGGACCCCAATCTGGACTCTTCCCGCCTTGGGA 1257
 Db 217 AAATGTTATCACCTTGTAGGGGACCCCAATCTGGACTCTTCCCGCCTTGGGA 158
 Gy 1258 CATCGAGCGCGGAGCAGTGGCCGCGAGCCCTGGGCGAGGAGTCTCAGGAGGCGCA 1317
 Db 157 CATCGAGCGCGGAGCAGTGGCCGCGAGCCCTGGGCGAGGAGTCTCAGGAGGCGCA 98
 Gy 1318 CTGAGCGCTGCTGGCGGAGCCCTGGGACATCCGAGCCACCGAGCAAGTCTCTGGGG 1377
 Db 97 CTGAGCGCTGCTGGCGGAGCCCTGGGACATCCGAGCCACCGAGCAAGTCTCTGGGG 38
 Gy 1378 CGACTCTGAATAAACCTTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1414
 Db 37 CGACTCTGAATAAACCTTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1

RESULT 12
 BC029819/c
 LOCUS Homo sapiens hypothetical protein FLJ32334, mRNA (cDNA clone
 DEFINITION MGC:35310 IMAGE:5177619), complete cds.
 ACCESSION BC029819
 VERSION BC029819.1 GI:20987583
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1923)
 AUTHORS Strausberg, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D., Klausner, R.D., Collins, F.S., Wagner, L., Shenman, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,

Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.B., Brownstein, M.J., Usdin, T.B., Tschuyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, W.A. (2002) Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22398257
 12477932
 2 (bases 1 to 1923)
 Strausberg, R.
 Direct Submission
 Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC Help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: angbcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loussegod, H., Kowis, C.R., Speed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 51 Row: 0 Column: 3
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21389332.

FEATURES
 source
 1..1923
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:35310 IMAGE:5177619"
 /tissue_type="Brain, Lung, Testis, adult, pooled whole"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..1923
 /genes="FLJ32334"
 /db_xref="LocusID:90527"
 401..1852
 /codon_start=1
 /product="FLJ32334 protein"
 /protein_id="AAH29819.1"
 /db_xref="GI:20987584"
 /db_xref="LocusID:90527"
 /translation="MATLGFTEFFVAGPKPTFPMDTTLASIMIFLTAATFVILPG
 IRKTLFLWLVVTSFLTGALIAVFSMSVQSTNTSKAFSSWISADIGLQ
 VGGVNTLTGTPOQLNATINVEFTWRLGNEAEVAKELGPPVLYLAEK
 FTRPGCYLQYRVLGHTSAMLVAFCLWLANVLSMPVAVGYMLATGIFOL
 LALLFPMATSLRSPCLHIGASVLTHTGHPAFWILTGLLCVLGLAMAVHROP
 HLKAFNDSVDEDMRLWSPEEGLLSPRYRSMADSPKQDIPLEASSTKAYRPR
 RLSPVADVRGLAPALSPGALLQAWRALPLGRCPKAGESLGGPHSPWRFP
 EGCERWAETGDSPLRGRGTGRLWRNGSKERRACGVRAMLPLVNSGLKRPSCIL

DLPCWYDRDARAFPHLEPTPCVTSHTPLI"

ORIGIN

Query Match 11.5%; Score 166; DB 9; Length 1923;
 Best Local Similarity 100.0%; Pred. No. 2.7e-80;
 Matches 166; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1216 CTGTGAGGGGACCAATCTGACCTCTCCCGCTTGGGACATCGCAGCGCGGAGC 1275
 DB 1557 CTGTGAGGGGACCAATCTGACCTCTCCCGCTTGGGACATCGCAGCGCGGAGC 1498
 QY 1276 AGTGCCCGCCAGGCTTGGGCGAGAGCTCCAGGAGGCACTGAGCGCTGCTGGCGG 1335
 DB 1497 AGTGCCCGCCAGGCTTGGGCGAGAGCTCCAGGAGGCACTGAGCGCTGCTGGCGG 1438
 QY 1336 AGGCTCTGGACATCCGACGACGAGGAGAGTCTCTGGGGCGAT 1381
 DB 1437 AGGCTCTGGACATCCGACGACGAGGAGAGTCTCTGGGGCGAT 1392

HS183A12R 218 bp DNA linear PRI 19-OCT-1995
 H.sapiens CpG island DNA genomic MseI fragment, clone 183a12,
 reverse read cpG183a12.rta.
 Z59954
 Z59954.1 GI:1031867
 CpG island; genomic MseI fragment.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 CROSS, S.H., Charlton, J.A., Nan, X. and Bird, A.P.
 Purification of CpG islands using a methylated DNA binding column
 Nat. Genet. 6 (3), 236-244 (1994)

94282070

8012384

REFERENCE 2 (bases 1 to 218)
 DODSWORTH, S.J., Huckle, E., Wilkinson, P. and Micklem, G.

Direct Submission
 Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1RO, England. E-mail contact: humquerry@sanger.ac.uk

Vector: pGEM-5zf(-)
 Clones are available from the UK MRC Human Genome Mapping Project
 Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
<http://www.hgmp.mrc.ac.uk/> for details
 or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES

Location/Qualifiers

1..218
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="183a12"
 /sex="male"
 /tissue_type="blood"
 /clone_lib="CGI-1"
 /dev_stage="adult"

ORIGIN

Query Match 9.4%; Score 136; DB 9; Length 218;
 Best Local Similarity 100.0%; Pred. No. 1.2e-63;
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1268 CGGGAGCAGTGGCGGCGCTGGGCGAGGAGCTCCAGGAGGCACTGAGCGCTG 1327
 DB 72 CGGGAGCAGTGGCGGCGCTGGGCGAGGAGCTCCAGGAGGCACTGAGCGCTG 131
 QY 1328 CTGGCGGAGGCTCGGACATCCGAGGACCGAGGAAAGTCTCTGGGGCGATCTGTAA 1387
 DB 132 CTGGCGGAGGCTCGGACATCCGAGGACCGAGGAAAGTCTCTGGGGCGATCTGTAA 191
 QY 1388 ATAAACCTTTTCT 1403

|||||
 192 ATAAACCTTTTCT 207

RESULT 14
 BC004161
 LOCUS

DEFINITION
 Homo sapiens transmembrane 4 superfamily member tetraspan NET-7,
 mRNA (cDNA clone MGC:2447 IMAGE:2958221), complete cds.
 BC004161
 BC004161.2 GI:33872773

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS

1 (bases 1 to 1725)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Sanchez, A., Whitling, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
 Scherzer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 1725)
 Strausberg, R.
 Direct Submission
 Submitted (01-MAR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:13278776.
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/BTP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadansystemsbio.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clon distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>

Series: IRAL Plate: 2 Row: k Column: 4

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 21264576.

Location/Qualifiers

1..1725

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:2447 IMAGE:2958221"

/tissue_type="kidney, renal cell adenocarcinoma"

/clone_lib="NIH_MGC_14"

FEATURES

source

```

/lab_host="DH10B-R"
/notes="vector: pOTB7"
1. .1725
/genes="NET-7"
/db_xref="LocusID:23555"
120. .1004
/codon_start=1
/product="tetraspan NET-7"
/protein_id="AA04161.1"
/db_xref="GI:13278777"
/db_xref="LocusID:23555"
/translation="MPRGSEQVRCARFSLYMLKFSLIYSTVFWLIGALVLSVGIY
AEVERQKXLTSAFLAPAILILGLVWFMVSPFGLVSLRDLNLYLQAFMYILGIC
LIMELIGGVVALTFNQTIDFLNDNRIGLNIYDLDLDFKNTMDVQKFKKCGGSDY
RWSKNQYHDCSAPGLACGVPTCCIRNTTEVNTMCGYKTIIDKERSVQDVIVYRG
CTNAVITWMDNTYIMAGILLGILLPQFLIGVLLTLYITRVEDIMEHSVTDLGLGPG
AKPSVEAAGTGCCCLCPN"
misc_feature
174. .854
/notes="transmembrane4; Region: Tetraspanin family"
/db_xref="CDD:pfam00335"

ORIGIN
Query Match 2.7%; Score 39; DB 9; Length 1725;
Best Local Similarity 100.0%; Pred. No. 7.8e-10;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1403 TTTTGTGTTTTTAAAAA...AAAAAAAAAAAAAAAAAAAAA 1441
Db 1683 TTTTGTGTTTTTAAAAA...AAAAAAAAAAAAAAAAAAAAA 1721

RESULT 15
BC003157 Homo sapiens transmembrane 4 superfamily member tetraspan NET-7,
LOCUS mRNA (cDNA clone MGC:4120 IMAGE:2958221), complete cds.
DEFINITION BC003157 1726 bp mRNA linear PRI 03-OCT-2003
VERSION BC003157.1 GI:13111968
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1726)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L.,
Schaeetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carroll, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Pahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalski, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
MEDLINE 12477932
REFERENCE 2 (bases 1 to 1726)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

```

```

USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Boudet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 9 Row: n Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 21264576.

FEATURES
Location/Qualifiers
1. .1726
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:4120 IMAGE:2958221"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_lib="NIH MGC 14"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1. .1726
/genes="NET-7"
/db_xref="LocusID:23555"
120. .1004
/codon_start=1
/product="tetraspan NET-7"
/protein_id="AAH03157.1"
/db_xref="GI:13111969"
/db_xref="LocusID:23555"
/translation="MPRGSEQVRCARFSLYMLKFSLIYSTVFWLIGALVLSVGIY
AEVERQKXLTSAFLAPAILILGLVWFMVSPFGLVSLRDLNLYLQAFMYILGIC
LIMELIGGVVALTFNQTIDFLNDNRIGLNIYDLDLDFKNTMDVQKFKKCGGSDY
RWSKNQYHDCSAPGLACGVPTCCIRNTTEVNTMCGYKTIIDKERSVQDVIVYRG
CTNAVITWMDNTYIMAGILLGILLPQFLIGVLLTLYITRVEDIMEHSVTDLGLGPG
AKPSVEAAGTGCCCLCPN"
misc_feature
180. .896
/notes="transmembrane4; Region: Tetraspanin family"
/db_xref="CDD:pfam00335"

ORIGIN
Query Match 2.7%; Score 39; DB 9; Length 1726;
Best Local Similarity 100.0%; Pred. No. 7.8e-10;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1403 TTTTGTGTTTTTAAAAA...AAAAAAAAAAAAAAAAAAAAA 1441
Db 1683 TTTTGTGTTTTTAAAAA...AAAAAAAAAAAAAAAAAAAAA 1721

Search completed: June 3, 2004, 15:07:01
Job time : 5815 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 3, 2004, 13:05:55 ; Search time 4062 Seconds
(without alignments)
10593.649 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441
Sequence: 1 aaagtaacggtacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0
Searched: 27513289 seqs, 14931090276 residues

Word size : 0
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: em_gss_hum:*
- 18: em_gss_inv:*
- 19: em_gss_pin:*
- 20: em_gss_vit:*
- 21: em_gss_fun:*
- 22: em_gss_mam:*
- 23: em_gss_mus:*
- 24: em_gss_pro:*
- 25: em_gss_red:*
- 26: em_gss_pug:*
- 27: em_gss_vri:*
- 28: gb_gss1:*
- 29: gb_gss2:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| C 1 | 528 | 36.6 | 537 | 9 AI821606 | AI821606 nk08a11.x |
| C 2 | 515 | 35.7 | 515 | 9 AI826629 | AI826629 wk35804.x |
| C 3 | 510 | 35.4 | 510 | 9 AI304327 | AI304327 q057n06.x |
| C 4 | 446 | 31.0 | 446 | 9 AI984141 | AI984141 wu21c02.x |

| | | | | | |
|------|-----|------|-----|----------------|---------------------|
| C 5 | 434 | 30.1 | 434 | 9 AI991272 | AI991272 wu41h04.x |
| C 6 | 414 | 28.7 | 414 | 9 AA593860 | AA593860 nn19f03.s |
| C 7 | 405 | 28.1 | 405 | 9 AI274929 | AI274929 q149c11.x |
| C 8 | 389 | 27.0 | 389 | 9 AI791844 | AI791844 nk08a11.y |
| C 9 | 388 | 26.9 | 388 | 9 AI660493 | AI660493 wu87h02.x |
| C 10 | 382 | 26.5 | 382 | 9 AI983793 | AI983793 wu20c09.x |
| C 11 | 381 | 26.4 | 381 | 9 AI281211 | AI281211 qk58e08.x |
| C 12 | 375 | 26.0 | 375 | 14 CB854121 | CB854121 UI-CF-DU1 |
| C 13 | 374 | 26.0 | 374 | 9 AI346155 | AI346155 qp43f12.x |
| C 14 | 369 | 25.6 | 369 | 9 AI660560 | AI660560 wu8b08.x |
| C 15 | 348 | 24.1 | 348 | 9 CB305399 | CB305399 UI-CF-EN1 |
| C 16 | 330 | 22.9 | 330 | 9 AA573825 | AA573825 nk08a11.s |
| C 17 | 317 | 22.0 | 317 | 12 BM987789 | BM987789 UI-H-COO- |
| C 18 | 296 | 20.5 | 296 | 9 AI821178 | AI821178 ne17a03.y |
| C 19 | 291 | 20.2 | 291 | 9 AW050605 | AW050605 w219b11.x |
| C 20 | 280 | 19.4 | 280 | 12 BM977010 | BM977010 UI-CF-EN1 |
| C 21 | 277 | 19.2 | 277 | 9 AI732165 | AI732165 ne17a03.x |
| C 22 | 270 | 18.7 | 270 | 14 CB850544 | CB850544 UI-CF-EN1 |
| C 23 | 260 | 18.0 | 260 | 9 AW009962 | AW009962 wu89h06.x |
| C 24 | 252 | 17.5 | 252 | 9 AI262416 | AI262416 qk38e04.x |
| C 25 | 185 | 12.8 | 185 | 539 9 AI924216 | AI924216 wn03b10.x |
| C 26 | 183 | 12.7 | 183 | 12 BM973444 | BM973444 UI-CF-EC1 |
| C 27 | 166 | 11.5 | 166 | 315 9 AA618335 | AA618335 nq15g11.s |
| C 28 | 149 | 10.3 | 149 | 9 AA469031 | AA469031 ne17a03.s |
| C 29 | 94 | 6.5 | 602 | 12 BQ017315 | BQ017315 UI-H-DT1- |
| C 30 | 88 | 6.1 | 151 | 9 AI695625 | AI695625 wu50h04.x |
| C 31 | 88 | 6.1 | 378 | 10 BF917041 | BF917041 IL3-UT011 |
| C 32 | 44 | 3.1 | 119 | 9 AI921289 | AI921289 wo22h06.x |
| C 33 | 42 | 2.9 | 942 | 14 CF241985 | CF241985 AGNCNCOURT |
| C 34 | 41 | 2.8 | 279 | 9 AI270429 | AI270429 qu84h03.x |
| C 35 | 40 | 2.8 | 175 | 14 CF316906 | CF316906 HD--06-G1 |
| C 36 | 40 | 2.8 | 217 | 14 CF328864 | CF328864 NACL--03- |
| C 37 | 39 | 2.7 | 188 | 12 BM887036 | BM887036 sam33408. |
| C 38 | 39 | 2.7 | 225 | 9 AI932794 | AI932794 wo38g03.x |
| C 39 | 39 | 2.7 | 243 | 13 BQ394290 | BQ394290 NISC_Rg08 |
| C 40 | 39 | 2.7 | 248 | 9 AI537827 | AI537827 tp35h10.x |
| C 41 | 39 | 2.7 | 250 | 14 CD421839 | CD421839 laa66d09. |
| C 42 | 39 | 2.7 | 284 | 9 AI537819 | AI537819 tp35g10.x |
| C 43 | 39 | 2.7 | 300 | 9 AI475331 | AI475331 tl81h09.x |
| C 44 | 39 | 2.7 | 324 | 10 AW130187 | AW130187 xf29f09.x |
| C 45 | 39 | 2.7 | 328 | 13 BU055639 | BU055639 UI-M-F00- |

ALIGNMENTS

RESULT 1
AI821606/c
LOCUS nk08a11.x5 NCI CGAP Co2 Homo sapiens cDNA clone IMAGE:1012892 3,
DEFINITION similar to contains TARI.t3 TARI MER22 repetitive element ;, mRNA
sequence.
ACCESSION AI821606
VERSION AI821606.1 GI:5440685
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 537)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other ESTs: nk08a11.y5
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing By: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image.html

This read is a RESEQUENCE of a previously sequenced human clone Original clone citation: National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
This read has been verified (found to hit its original self in the correct orientation)

Insert Length: 741 Std Error: 0.00
Seq primer: -40UP from Gibco

High quality sequence stop: 459.

FEATURES

Location/Qualifiers
1..537
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1012892"
/tissue_type="tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP Co2"
/note="Organ: colon; Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: oligo dT. Bulk colon villous adenoma. 5' adaptor sequence: 5' GAATTCGGCAGAG 3' 3' adaptor sequence: 5' CTCGAGTTTCTTTT 3' Average insert size: 1.1 kb."

ORIGIN

Query Match 36.6%; Score 528; DB 9; Length 537;
Best Local Similarity 100.0%; Pred. No. 1.6e-95;
Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 878 GCTGGCACTGTGACCAACGAGGCTTCGGGCTCTTCGGGGTCTTCGGCTTCGCA 937
Db 537 GCTGGCACTGTGACCAACGAGGCTTCGGGCTCTTCGGGGTCTTCGGCTTCGCA 478
Qy 938 TCTTAGCGTGGCGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCG 997
Db 477 TCTTAGCGTGGCGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCG 418
Qy 998 GCGCGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCG 1057
Db 417 GCGCGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCG 358
Qy 1058 TGGTAGTCTTCAGTATGTTGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCG 1117
Db 357 TGGTAGTCTTCAGTATGTTGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCG 298
Qy 1118 AGGACTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1177
Db 297 AGGACTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 238
Qy 1178 AGGCGCTCTCCAGACTTAAATGATACCACTAACTGTGAGGGGAGCCCAATCTGG 1237
Db 237 AGGCGCTCTCCAGACTTAAATGATACCACTAACTGTGAGGGGAGCCCAATCTGG 178
Qy 1238 ACTCTTCCTCCGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGG 1297
Db 177 ACTCTTCCTCCGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGG 118
Qy 1298 GGAGAGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1357
Db 117 GGAGAGCTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 58
Qy 1358 CCAGGAGAAAGTCTCCCTGGGGGAGTCTGTAATAAATCTTTTCTTTTCTTTT 1405
Db 57 CCAGGAGAAAGTCTCCCTGGGGGAGTCTGTAATAAATCTTTTCTTTTCTTTT 10

RESULT 2
AI826629/c
LOCUS
DEFINITION
515 bp mRNA linear EST 21-DEC-1999
wk35e04.x1 NCI_CGAP_Pr22 Homo sapiens CDNA clone IMAGE:2417406 3',

mRNA sequence.

AI826629
AI826629.1 GI:5447300

EST.
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 515)

AUTHORS
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

JOURNAL
Tumor Gene Index

COMMENT
Unpublished (1997)

Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image.html

Insert Length: 910 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 467.

Location/Qualifiers
1..515

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2417406"

/sex="male"

/tissue_type="normal prostate"

/lab_host="DH10B"

/clone_lib="NCI_CGAP_Pr22"

/note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from normal prostate bulk tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 35.7%; Score 515; DB 9; Length 515;
Best Local Similarity 100.0%; Pred. No. 6.3e-93;
Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 902 CTTTCGGGCTCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTC 961
Db 515 CTTTCGGGCTCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTC 456
Qy 962 TCGGCTAGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGG 1021
Db 455 TCGGCTAGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTC 396
Qy 1022 CAACCGGCGTCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTC 1081
Db 395 CAACCGGCGTCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTC 336
Qy 1082 CCAGGCTCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGG 1141
Db 335 CCAGGCTCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGG 276
Qy 1142 GCTACCTCTTATCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTC 1201
Db 275 GCTACCTCTTATCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTCGGGCTTC 216
Qy 1202 GTATCACCACTAACTGTGAGGGGAGCCCAATCTTGGACTCTTCCCGCTTCGGGCTTC 1261
Db 215 GTATCACCACTAACTGTGAGGGGAGCCCAATCTTGGACTCTTCCCGCTTCGGGCTTC 156

| | | | |
|--|------|---|------|
| QY | 1262 | GCAGGCGGGAAGCAGTGGCCGACGCTTGGCCAGGAGAGTCCACGAGAGGGCACTGA | 1322 |
| Db | 155 | GCAGGCGGGAAGCAGTGGCCGACGCTTGGCCAGGAGAGTCCACGAGAGGGCACTGA | 96 |
| QY | 1322 | CGCGTGTGCGCGAGCGCTTCGGACATCCGACAGGCACACAGGAGAAAGTCTCTCTGGGGCGAT | 1381 |
| Db | 95 | CGCGTGTGCGCGAGCGCTTCGGACATCCGACAGGCACACAGGAGAAAGTCTCTCTGGGGCGAT | 36 |
| QY | 1382 | CTGTAAATAAACCTTTTCTTTCTTTGTTTTTAA | 1416 |
| Db | 35 | CTGTAAATAAACCTTTTCTTTCTTTGTTTTTAA | 1 |
| RESULT 3 | | | |
| AI304327/c | | | |
| LOCUS | | | |
| DEFINITION | | | |
| Q057906.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1912667 3' | | | |
| similar to contains TARI.t3 TARI repetitive element ;, mRNA | | | |
| sequence. | | | |
| ACCESSION | | | |
| VERSION | | | |
| SOURCE | | | |
| KEYWORDS | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| COMMENT | | | |
| AI304327 | | | |
| Q057906.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1912667 3' | | | |
| similar to contains TARI.t3 TARI repetitive element ;, mRNA | | | |
| sequence. | | | |
| ACCESSION | | | |
| VERSION | | | |
| SOURCE | | | |
| KEYWORDS | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| COMMENT | | | |
| AI304327 | | | |
| Q057906.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1912667 3' | | | |
| similar to contains TARI.t3 TARI repetitive element ;, mRNA | | | |
| sequence. | | | |
| ACCESSION | | | |
| VERSION | | | |
| SOURCE | | | |
| KEYWORDS | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| COMMENT | | | |
| AI304327 | | | |
| Q057906.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1912667 3' | | | |
| similar to contains TARI.t3 TARI repetitive element ;, mRNA | | | |
| sequence. | | | |
| ACCESSION | | | |
| VERSION | | | |
| SOURCE | | | |
| KEYWORDS | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| COMMENT | | | |
| AI304327 | | | |
| Q057906.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1912667 3' | | | |
| similar to contains TARI.t3 TARI repetitive element ;, mRNA | | | |
| sequence. | | | |
| ACCESSION | | | |
| VERSION | | | |
| SOURCE | | | |
| KEYWORDS | | | |
| ORGANISM | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| COMMENT | | | |
| AI304327 | | | |
| Q057906.x1 NCI_CGAP_Co8 Homo sapiens cDNA clone IMAGE:1912667 3' | | | |
| similar to contains TARI.t3 TARI repetitive element ;, mRNA | | | |
| sequence. | | | |

dieck@im.wustl.edu); colonic mucosa represents a range of disease involvement from moderate to severe. Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Ronaldo.

ORIGIN

| Query Match | 31.08; | Score 446; | DB 9; | Length 446; |
|-----------------------|-----------------|--|-----------|-------------|
| Best Local Similarity | 100.0%; | Pred. No. 3e-79; | | |
| Matches 446; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 971 | GCTCTCCGCGCTCACCACTCAGTATACGGCGCGCGCTTCTGGGTGCACGCTGGCAACCGGCG | 1030 | |
| Db | 446 | GCTCTCCGCGCTCACCACTCAGTATACGGCGCGCGCTTCTGGGTGCACGCTGGCAACCGGCG | 387 | |
| Qy | 1031 | TCTGTGTGCTTCTCTCGGAGGGGCGGTGTGTAGTCTCCAGTATGTTTCGGGCCACAGGCTC | 1090 | |
| Db | 386 | TCTGTGTGCTTCTCTCGGAGGGGCGGTGTGTAGTCTCCAGTATGTTTCGGGCCACAGGCTC | 327 | |
| Qy | 1091 | TTGCGACCCCTTCTGGACCAAGGCCCAAGCACTGTCAGCCAGGAGAGAGGGGGCTCACCTC | 1150 | |
| Db | 326 | TTGCGACCCCTTCTGGACCAAGGCCCAAGCACTGTCAGCCAGGAGAGAGGGGGCTCACCTC | 267 | |
| Qy | 1151 | TTATCTCTGGGAGCCCACTGCAAGAAGAGGGCGCTCTCCAGACTTTAAATGTATCAACA | 1210 | |
| Db | 266 | TTATCTCTGGGAGCCCACTGCAAGAAGAGGGCGCTCTCCAGACTTTAAATGTATCAACA | 207 | |
| Qy | 1211 | CTAACTGTGAGGGGGACCAATCTGGACTCTCTCCCGCGCTTGGGACATCGCAGAGGCGG | 1270 | |
| Db | 206 | CTAACTGTGAGGGGGACCAATCTGGACTCTCTCCCGCGCTTGGGACATCGCAGAGGCGG | 147 | |
| Qy | 1271 | GAAGCATGCCC CGCCAGGCTTGGGCCCAAGGAGAGCTCCAGGAGGGCACTGAGCGCTGCTG | 1330 | |
| Db | 146 | GAAGCATGCCC CGCCAGGCTTGGGCCCAAGGAGAGCTCCAGGAGGGCACTGAGCGCTGCTG | 87 | |
| Qy | 1331 | GCGCGAGGCTCTGAGACATCGCAGGCAACCGGAAAGTCTCTCTGGGCGGATCTGTAAATA | 1390 | |
| Db | 86 | GCGCGAGGCTCTGAGACATCGCAGGCAACCGGAAAGTCTCTCTGGGCGGATCTGTAAATA | 27 | |
| Qy | 1391 | AACCTTTTTTCTTTCTTTTAAAA | 1416 | |
| Db | 26 | AACCTTTTTTCTTTCTTTTAAAA | 1 | |

| | | | | | | |
|------------|------------|--|--------|------|--------|-----------------|
| RESULT 5 | AI991272/c | AI991272 | 434 bp | mrna | linear | EST 09-MAR-2000 |
| LOCUS | | ws41h04.x1 | | | | |
| DEFINITION | | Soares Dieckgraefe colon_NHCD Homo sapiens cDNA clone IMAGE:2522647 3'. mRNA sequence. | | | | |

| | | |
|-----------|--|------------|
| ACCESSION | AI991272 | |
| VERSION | AI991272.1 | GI:5838177 |
| KEYWORDS | EST. | |
| SOURCE | Homo sapiens (human) | |
| ORGANISM | Homo sapiens | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| | 1. (bases 1 to 434) | |
| REFERENCE | NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap . | |
| AUTHORS | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), | |
| TITLE | Tumor Gene Index | |
| JOURNAL | Unpublished (1997) | |

unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through iLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 558 Std Error: 0.00
Seq primer: -40UP from Gibco.

```

FEATURES
source
Location/Qualifiers
1. .434
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="IMAGE:2522647"

```

/tissue_type="colonic mucosa from 3 patients with Crohn's disease"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares Dieckgraefe_colon_NHCD"
/note="organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGAAGTCGGACGCGCCGCTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraefe (Washington University, dieck@im.wustl.edu); colonic mucosa represents a range of disease involvement from moderate to severe Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

| Query Match | 30.1% | Score 434; | DB 9; | Length 434; |
|-----------------------|-----------------|---|-----------|-------------|
| Best Local Similarity | 100.0%; | Prod. No. 7.2e-77; | | |
| Matches 434; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 981 | GCTCACACATCAGTACGGCGCGCCTCTCTGGGTCAAGCTGGGCACACGGCGCTCTCTGTGCCT | 1040 | |
| Db | 434 | GCTCACACATCAGTACGGCGCGCCTCTCTGGGTCAAGCTGGGCACACGGCGCTCTCTGTGCCT | 375 | |
| Qy | 1041 | CTTCTCTCGAGAGGGCGGTGTGAGTCTCCAGTATGTTTCGGCCCAAGCGCTCTTCGCAACCT | 1100 | |
| Db | 374 | CTTCTCTCGAGAGGGCGGTGTGAGTCTCCAGTATGTTTCGGCCCAAGCGCTCTTCGCAACCT | 315 | |
| Qy | 1101 | TCGTGGACCAAGCGCCAAAGGACATGCAGCAGGAGAGAGGGGGCTCACCTCTTATCTCTCGG | 1160 | |
| Db | 314 | TCGTGGACCAAGCGCCAAAGGACATGCAGCAGGAGAGAGGGGGCTCACCTCTTATCTCTCGG | 255 | |
| Qy | 1161 | CGACCCACTGCAACAGCAGGCGCGCTCTCCAGACTTTAAATATGATACCACTAACCTGTG | 1220 | |
| Db | 254 | CGACCCACTGCAACAGCAGGCGCGCTCTCCAGACTTTAAATATGATACCACTAACCTGTG | 195 | |
| Qy | 1221 | AGGGGACCCCAATCTGGACTCTTTCCCGGCTTTGGGACATCGCAGGCGCGGGAAGCAGTGC | 1280 | |
| Db | 194 | AGGGGACCCCAATCTGGACTCTTTCCCGGCTTTGGGACATCGCAGGCGCGGGAAGCAGTGC | 135 | |
| Qy | 1281 | CCGCGCAGGCTGTGGCCAGGAGAGCTCCAGAAAGGCACTGAGGCGCTCTGGCCGCGAGGCC | 1340 | |
| Db | 134 | CCGCGCAGGCTGTGGCCAGGAGAGCTCCAGAAAGGCACTGAGGCGCTCTGGCCGCGAGGCC | 75 | |
| Qy | 1341 | TCGGACATCCGAGGACACCAAGGAAAGTCTCTCGGGCGGATCTGTAAATAAAACCTTTTTT | 1400 | |
| Db | 74 | TCGGACATCCGAGGACACCAAGGAAAGTCTCTCGGGCGGATCTGTAAATAAAACCTTTTTT | 15 | |
| Qy | 1401 | TCCTTTGTGTTTTTA 1414 | | |
| Db | 14 | TCCTTTGTGTTTTTA 1 | | |

| | | | | | |
|------------|--|---------------|--------------|------------------------------|-----------------|
| RESULT 6 | AA593860/c | 414 bp | mRNA | linear | EST 25-SEP-1997 |
| LOCUS | nm19f03.s1 | NCI_CGAP_Col2 | Homo sapiens | cDNA clone IMAGE:1084349 3', | |
| DEFINITION | mRNA sequence. | | | | |
| ACCESSION | AA593860 | | | | |
| VERSION | AA593860.1 | GI:2408538 | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |
| | 1 (bases 1 to 414) | | | | |
| REFERENCE | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . | | | | |
| AUTHORS | National Cancer Institute, Cancer Genome Anatomy Project (CGAP). | | | | |
| TITLE | Tumor Gene Index | | | | |

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgepbs-remail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. Emert-Buck, M.D., Ph.D.

DNA Library Preparation: Stratagene, Inc.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnlnl.gov/bbrp/image/image.html
Insert Length: 1204 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 406.
Location/Qualifiers
1. .414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1084349"
/sex="mixed"
/tissue_type="colon tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP C012"
/note=Organ: colon; Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: oligo dt. Pooled colon tumors. 5' adaptor sequence: 5'-GAATTCGCACAG 3' adaptor sequence: 5'-CTCAGATTGTGTGTGTGTGTGTGTGT 3' Average insert size: 1.2 kb.

| Query Match | 28.7%; Score 414; DB 9; Length 414; |
|--|--|
| Best Local Similarity | 100.0%; Pred. No. 6.7e-73; |
| Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| QY | CGCGCGCTTCTGGGTCAAGCTGGCAACCGAGCGCTCTGTGCTCTCTCTCTCGAGGGGCGGT 1058 |
| DB | |
| 414 | CGCGCGCTTCTGGGTCAAGCTGGCAACCGAGCGCTCTGTGCTCTCTCTCTCGAGGGGCGGT 355 |
| QY | GGTGAGTCTCCAGTATGTTGGGCCCAAGCGCTCTTGGCACCCCTCTGGACCAAAAGCGCCAA 1158 |
| DB | |
| 354 | GGTGAGTCTCCAGTATGTTGGGCCCAAGCGCTCTTGGCACCCCTCTGGACCAAAAGCGCCAA 295 |
| QY | GGATCTGCAGCACGAGGAGGGGGCTCACTCTTATCTCTGGGACCCACTGTCACAGCA 1178 |
| DB | |
| 294 | GGATCTGCAGCACGAGGAGGGGGCTCACTCTTATCTCTGGGACCCACTGTCACAGCA 235 |
| QY | GGCGGCTCTCCAGACTTAAATGTATACCACTAACTGTGAGGGGGAACCAATCTGGA 1238 |
| DB | |
| 234 | GGCGGCTCTCCAGACTTAAATGTATACCACTAACTGTGAGGGGGAACCAATCTGGA 175 |
| QY | CTCCTTCCGCGGCTTGGGACATCGGAGGCCGGGAGAGCATGTCGCCGCGAGGCTGGGGCAG 1298 |
| DB | |
| 174 | CTCCTTCCGCGGCTTGGGACATCGGAGGCCGGGAGAGCATGTCGCCGCGAGGCTGGGGCAG 115 |
| QY | GAGAGCTTCCAGGAAGGGCACTGAGCGCTCTGTCGCCGAGGCGCTCGGACATCTCGCAGGCAC 1358 |
| DB | |
| 114 | GAGAGCTTCCAGGAAGGGCACTGAGCGCTCTGTCGCCGAGGCGCTCGGACATCTCGCAGGCAC 55 |
| QY | CAGGAAAAGTCTCTCGGGGGGATCTGTAAATAAACCTTTTTTCTTTTGTGTTTTT 1412 |
| DB | |
| 54 | CAGGAAAAGTCTCTCGGGGGGATCTGTAAATAAACCTTTTTTCTTTTGTGTTTTT 1 |

A1274929/c
 LOCUS A1274929 405 bp mRNA linear EST 29-JAN-1999
 Q14911.x1 NCI CGAP C08 Homo sapiens cDNA clone IMAGE:1875668 3,
 similar to contains TAR1.t3 TAR1 repetitive element, mRNA
 sequence.
 A1274929
 accession A1274929.1 GI:3897203
 version
 keywords EST.

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 405)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/nciccap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg. Ph.D.

Contact: Robert Strausberg, Ph.D.
 Email: cgaps@emil.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lemmon, Ph.D.
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert length: 1458 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 395.

```
location/Qualifiers
1. 405
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="IMAGE:1875658"
/tissue_type="adenocarcinoma"
/lab_host="PH103"
/seq_lib_name="seq_088"
```

| Query Match | 28.1% | Score 405; | DB 9; | Length 405; |
|-----------------------|---|--------------------|-----------|-------------|
| Best Local Similarity | 100.0%; | Pred. No. 4.1e-71; | | |
| Matches 405; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| 1009 | TGGGTCAACCTCGCAACCGGGCTCTGTGCTCTTCTCGAGGGGCGGTGGTGAAGTCTC | 1068 | | |
| 405 | TGGGTCAACCTCGCAACCGGGCTCTGTGCTCTTCTCGAGGGGCGGTGGTGAAGTCTC | 346 | | |
| 1069 | CAGTATGTTGGGCCCGACGGCTCTTTGCGACCCCTTCTGGACCAAAAGCGCAAGGACTGCAGC | 1128 | | |
| 345 | CAGTATGTTGGGCCCGACGGCTCTTTGCGACCCCTTCTGGACCAAAAGCGCAAGGACTGCAGC | 286 | | |
| 1129 | CAGGAGAGGGGGGTCACTCTTATCTCTGGGCGACCACTGCACAGCAGGCGGCTCTC | 1188 | | |
| 285 | CAGGAGAGGGGGGTCACTCTTATCTCTGGGCGACCACTGCACAGCAGGCGGCTCTC | 226 | | |
| 1189 | CCAGACTTAAATGTATCACCACTAACTGTGAGGGGACCCCAATCTGGAAGTCTCTTCCCC | 1248 | | |
| 225 | CCAGACTTAAATGTATCACCACTAACTGTGAGGGGACCCCAATCTGGAAGTCTCTTCCCC | 166 | | |
| 1249 | GCCTTGGGACATCGCAGCGCGGGGAAGCAATGTCGCCCGCAGGCTGTGGGCAGGAGAGTCCA | 1308 | | |
| 165 | GCCTTGGGACATCGCAGCGCGGGGAAGCAATGTCGCCCGCAGGCTGTGGGCAGGAGAGTCCA | 106 | | |
| 1309 | GGAAAGGGCACTCAGACGGCTGTCTGGCGCGAGGGCTCTGGGACATCCGCGAGGCACACAGGAAAGT | 1368 | | |
| 105 | GGAAAGGGCACTCAGACGGCTGTCTGGCGCGAGGGCTCTGGGACATCCGCGAGGCACACAGGAAAGT | 46 | | |
| 1369 | CTCTCTGGGGCGGATCTGTAAATAAAACCTTTTTTCTTTTGTTTTTT | 1413 | | |
| 45 | CTCTCTGGGGCGGATCTGTGTAAATAAAACCTTTTTTCTTTTGTTTTTT | 1 | | |

916 GGGGTCTTTCGGCTTGGCTCCATCTCTAGCGTCCGGCTCTGCGCCGGTCCGCCCTAGGCTCC 975
181 GGGGTCTTTCGGCTTGGCTCCATCTCTAGCGTCCGGCTCTGCGCCGGTCCGCCCTAGGCTCC 240
976 TCGCGGTCTACCACTCAGTACGCGCGCGCTCTGGGTACGCTGGCAACCGCGGCTCTG 1035
241 TCGCGGTCTACCACTCAGTACGCGCGCGCTCTGGGTACGCTGGCAACCGCGGCTCTG 300
1036 TGCCTCTTCTCGAGGGGCGCTGGTGGTCTCCAGTATGTTCCGCGCCAGCGCTCTTCGC 1095
301 TGCTCTTCTCGAGGGGCGCTGGTGGTCTCCAGTATGTTCCGCGCCAGCGCTCTTCGC 360
1096 ACCCTTCTGGACCAAGCGCCAGGACTG 1124
361 ACCCTTCTGGACCAAGCGCCAGGACTG 389

RESULT 9
AI660493/c
LOCUS
DEFINITION
w667h02.x1 Soares Dieckgraeft colon_MHCD Homo sapiens cDNA clone
IMAGE:2346195.3 similar to contains TAR1.t3 TAR1 repetitive
element ; mRNA sequence.
ACCESSION
AI660493
VERSION
AI660493.1 GI:4764063
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 538)
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LInL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 676 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 463.
Location/Qualifiers
1. 538
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2346195"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares Dieckgraeft colon_MHCD"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not 1 - oligo(dT) primer [5',
TGTACCACTCTAGTGGGAGCGCGCGCTCTTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not 1 and cloned into the Not 1
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraeft (Washington University,
dieck@im.wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Soares and M. Fatima Bonaldo."

FEATURES
source
1. 538
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2346195"
/tissue_type="colonic mucosa from 3 patients with Crohn's
disease"
/lab_host="DH10B (phage-resistant)"
/clone_lib="Soares Dieckgraeft colon_MHCD"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not 1 - oligo(dT) primer [5',
TGTACCACTCTAGTGGGAGCGCGCGCTCTTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not 1 and cloned into the Not 1
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Tissue samples
provided by Dr. Brian Dieckgraeft (Washington University,
dieck@im.wustl.edu); colonic mucosa represents a range of
disease involvement from moderate to severe Crohn's
disease; samples include both perforating (fistulas) and
non-perforating samples. Library constructed by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 26.9%; Score 388; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 7.1e-68; Indels 0; Gaps 0;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1027 GCGGTCTTTCGGCTTGGCTCCATCTCTAGCGTCCGGCTCTGCGCCGGTCCGCCCTAGGCTCC 1086

AI791844 505 bp mRNA linear EST 13-DEC-1999
nk08all.y5 NCI_CGAP Co2 Homo sapiens cDNA clone IMAGE:1012892.5
similar to contains Alu repetitive element; mRNA sequence.
AI791844
VERSION
AI791844.1 GI:5339486
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 505)
CDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
JOURNAL
Unpublished (1997)
COMMENT
Other ESTs: nk08all.s1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
This read is a RESEQUENCE of a previously sequenced human clone
Original clone citation: see original entry for original citation
information
This 5' resequenced clone has no previous 5' data to verify this
new read against
Putative full length read
The vector to vector length is 671
Insert Length: 741 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 418.
Location/Qualifiers
1. 505
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1012892"
/tissue_type="tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI_CGAP Co2"
/note="Organ: colon; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dr. Bulk colon villous adenoma. 5' adaptor sequence:
5' GAATTCGGACGAG 3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."

ORIGIN
Query Match 27.0%; Score 389; DB 9; Length 505;
Best Local Similarity 100.0%; Pred. No. 4.8e-68; Indels 0; Gaps 0;
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
736 AAGTTCACACCGAGTACGCTTTCGGCTCTGACACAGTACCACTGGCGGACACTAC 795
1 AAGTTCACACCGAGTACGCTTTCGGCTCTGACACAGTACCACTGGCGGACACTAC 60
796 GCTTCGGCCACGATAGGCTGCTGCTCTGCTCTCTCCACGCTGCTCTCC 855
61 GCTTCGGCCACGATAGGCTGCTGCTCTCTCCACGCTGCTCTCC 120
856 ACGCGGCGCGCTCTACGAGGCGCTGGCACTGCTGACACCGAGCTTCGCGCTCTTC 915
121 ACGCGGCGCGCTCTACGAGGCGCTGGCACTGCTGACACCGAGCTTCGCGCTCTTC 180

```
1. .381
//organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
```

Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA sequence: 533-573, >ALU (matched complement)
 Seq primer: M13 FORWARD
 POLYA=No.

Location/Qualifiers
 1. 573
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="UI-CF-DUI-aal-1-16-0-UI"
 /tissue_type="Primary Lung Epithelial Cells"
 /dev_stage="Adult"
 /lab_host="DHI08 (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-CF-DUI"
 /note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-DUI is a normalized cDNA library containing the following tissue(s): Primary Lung Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCTGTAGGC.
 TAG_SEQ=None found"

FEATURES
 source

| ORIGIN | Query Match | Best Local Similarity | Score 375; DB 14; Length 573; | Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
|--------|-------------|--|-------------------------------|--|
| QY | 1044 | CCTCGAGGGGCGGTGTGAGTCTCCAGTATGTTTCGGCCGAGCGCTCTTCGACCCCTCT | 1103 | |
| Db | 482 | CCTCGAGGGGCGGTGTGAGTCTCCAGTATGTTTCGGCCGAGCGCTCTTCGACCCCTCT | 423 | |
| QY | 1104 | GGACCAAGCGCCCAAGGACTCGACGAGAGAGGGGGCTCACCTCTTATCTCGGCGA | 1163 | |
| Db | 422 | GGACCAAGCGCCCAAGGACTCGACGAGAGAGGGGGCTCACCTCTTATCTCGGCGA | 363 | |
| QY | 1164 | CCCACTGCACAGCAGCGCGCTCTCCAGACTTAAATGTATACCACTAACCTGTGAGG | 1223 | |
| Db | 362 | CCCACTGCACAGCAGCGCGCTCTCCAGACTTAAATGTATACCACTAACCTGTGAGG | 303 | |
| QY | 1224 | GGGACCCCAATCTGGACTCTTCCCGCTTGGACATCGCAGCGCGGAGCAGTGCCTG | 1283 | |
| Db | 302 | GGGACCCCAATCTGGACTCTTCCCGCTTGGACATCGCAGCGCGGAGCAGTGCCTG | 243 | |
| QY | 1284 | CCAGGCTTGGGCGCAGGAGAGCTCCAGGAAGGCGACTGAGCGCTGTGCGGAGGCTCG | 1343 | |
| Db | 242 | CCAGGCTTGGGCGCAGGAGAGCTCCAGGAAGGCGACTGAGCGCTGTGCGGAGGCTCG | 183 | |
| QY | 1344 | GACATCCGAGGCGCAGGAGAGCTCTCGGGCGATCTGTAAATAAACCTTTTCT | 1403 | |
| Db | 182 | GACATCCGAGGCGCAGGAGAGCTCTCGGGCGATCTGTAAATAAACCTTTTCT | 123 | |
| QY | 1404 | TTTGTGTTTTTAAAAA 1418 | | |
| Db | 122 | TTTGTGTTTTTAAAAA 108 | | |

RESULT 13
 AI346155/c
 LOCUS AI346155 405 bp mRNA linear EST 02-FEB-1999

/clone="IMAGE:1873190"
 /tissue_type="adenocarcinoma"
 /lab_host="DHI08"
 /clone_lib="NCI_CCAP_C08"
 /note="Organ: colon; Vector: p773-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

| ORIGIN | Query Match | Best Local Similarity | Score 381; DB 9; Length 381; | Matches 381; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
|--------|-------------|---|------------------------------|--|
| QY | 1033 | CTGTGCTCTTCTCTCGAGGGGCGGTGTGAGTCTCCAGTATGTTTCGGCCGAGCGCTCTT | 1092 | |
| Db | 381 | CTGTGCTCTTCTCTCGAGGGGCGGTGTGAGTCTCCAGTATGTTTCGGCCGAGCGCTCTT | 322 | |
| QY | 1093 | CGACCTTCTCGACCAAGCGCCAGGACTCGACGAGAGAGGGGGCTCACTCTT | 1152 | |
| Db | 321 | CGACCTTCTCGACCAAGCGCCAGGACTCGACGAGAGAGGGGGCTCACTCTT | 262 | |
| QY | 1153 | ATCCTCGGCGACCACTGCACAGCAGGCGGTCTCTCCAGACTTAAATGTATCACTACT | 1212 | |
| Db | 261 | ATCCTCGGCGACCACTGCACAGCAGGCGGTCTCTCCAGACTTAAATGTATCACTACT | 202 | |
| QY | 1213 | ACCTGTGAGGGGACCAATCTGACTCTTCCCGCTTGGACATCGCAGGCGGGA | 1272 | |
| Db | 201 | ACCTGTGAGGGGACCAATCTGACTCTTCCCGCTTGGACATCGCAGGCGGGA | 142 | |
| QY | 1273 | AGCAGTGCCTCGGCGCAGGCGGTCTCCAGAGAGCTCCAGGAAGGCGACTGAGCGCTGTGCG | 1332 | |
| Db | 141 | AGCAGTGCCTCGGCGCAGGCGGTCTCCAGAGAGCTCCAGGAAGGCGACTGAGCGCTGTGCG | 82 | |
| QY | 1333 | GGAGGCGCTCGGACATCGCAGGCGCAGGAGAAAGTCTCTCGGGCGATCTGTAAATAA | 1392 | |
| Db | 81 | GGAGGCGCTCGGACATCGCAGGCGCAGGAGAAAGTCTCTCGGGCGATCTGTAAATAA | 22 | |
| QY | 1393 | CTTTTCTTTTGTGTTTTT 1413 | | |
| Db | 21 | CTTTTCTTTTGTGTTTTT 1 | | |

RESULT 12
 CB854121/c
 LOCUS
 DEFINITION UI-CF-DUI-aal-1-16-0-UI.s1 UI-CF-DUI Homo sapiens cDNA clone
 UI-CF-DUI-aal-1-16-0-UI 3', mRNA sequence.
 CB854121
 VERSION CB854121.1 GI:30044498
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 573)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT
 Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171
 Email: paul-mccray@uiowa.edu

| | | |
|---------------------------|--|---------------------------------|
| 105 | GGAAGGCATGTAGCGCTGTGGCGGAGSCTTCGACATCGCAGGACCAACGGAAGT | 416 |
| 1369 | CTCCTGGGGCGATC 1382 | |
| 45 | CTCCTGGGGCGATC 32 | |
| RESULT 14 | | |
| AI660560/c | | |
| LOCUS | | |
| DEFINITION | | |
| ACCESSION | | |
| VERSION | | |
| KEYWORDS | | |
| SOURCE | | |
| ORGANISM | | |
| REFERENCE | | |
| AUTHORS | | |
| TITLE | | |
| JOURNAL | | |
| COMMENT | | |
| FEATURES | | |
| source | | |
| ORIGIN | | |
| Query Match | 25.6% | Score 369; DB 9; Length 486; |
| Best Local Similarity | 99.8% | Pred. NO. 4.4e-64; |
| Matches 419; Conservative | 0; | Mismatches 1; Indels 0; Gaps 0; |
| 1009 | TGGGTCAACCTGGACCGGCTCTGTGCTCTTCTCGAGGGGCGGTGAGTCTC | 1068 |
| 420 | TGGGTCAACCTGGACCGGCTCTGTGCTCTTCTCGAGGGGCGGTGAGTCTC | 361 |
| 1069 | CAGTATGTTTGGGCCACAGCGCTCTTTCGACCCCTTTCGACCAAAAGCGCAAGGACTGCAGC | 1128 |
| 360 | CAGTATGTTTGGGCCACAGCGCTCTTTCGACCCCTTTCGACCAAAAGCGCAAGGACTGCAGC | 301 |
| 1129 | CAGGAGAGAGGGGGCTACCTCTTATCTCTGGGCAACCACTGCACAGAGGGCGCTCTC | 1188 |
| 300 | CAGGAGAGAGGGGGCTACCTCTTATCTCTGGGCAACCACTGCACAGAGGGCGCTCTC | 241 |

1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT. 6hr to LPS 24h

ORIGIN

```

Query Match      24.1%; Score 348; DB 14; Length 737;
Best Local Similarity 100.0%; Pred.No. 4.le-60;
Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1027 GGGGTCCTGTGCGCTCTTCTCGAGGGGCGCTGTGGAGTCTCCAGTATGTTTCGGGCCAGC 1086
      |||||
353 GCGGTCTCTGTGCGCTCTTCTCGAGGGGCGCTGTGGAGTCTCCAGTATGTTTCGGGCCAGC 412
      |||||
1087 GCTCTTCGACCCCTTCTCGACCAAGGCCCAAGGACTGCAGCCAGGAGAGGGGCTCTCA 1146
      |||||
413 GCTCTTCGACCCCTTCTGGAACCAAGGCCCAAGGACTGCAGCCAGGAGAGGGGCTCTCA 472
      |||||
1147 CCTCTTATCTCGGCGACCCACTGCACAGCAGGCGCGCTCTCCAGACTTTAAATGTATC 1206
      |||||
473 CCTCTTATCTCGGCGACCCACTGCACAGCAGGCGCGCTCTCCAGACTTTAAATGTATC 532
      |||||
1207 ACCACTAACTGTGAGGGGGGCCCAACTTGAGCTCTTCCCGCTTGGGACATCCGAGG 1266
      |||||
533 ACCACTAACTGTGAGGGGGGCCCAACTTGAGCTCTTCCCGCTTGGGACATCCGAGG 592
      |||||
1267 CCGGGAACAGTGCCTCCGACGCTTGGGCGCAGGAGACTCCAGGAAGGCACTGAGCGCT 1326
      |||||
593 CCGGGAACAGTGCCTCCGACGCTTGGGCGCAGGAGACTCCAGGAAGGCACTGAGCGCT 652
      |||||
1327 GCTGCGCGAGCGCTCGACATCCCGAGCCACCGAGGAAAGTCTCCTG 1374
      |||||
653 GCTGCGCGAGCGCTCGACATCCCGAGCCACCGAGGAAAGTCTCCTG 700
      |||||

Search completed: June 3, 2004, 16:14:54
Job time : 4067 secs

```

Search completed: June 3, 2004, 16:14:54
Job time : 4067 secs

| | | |
|------|---|------|
| 1189 | CGAGACTTAAATATGATACACCACTTAACCTGTGGGGGGGACCCCAATCTGGACTCCTTCCCC | 1248 |
| 1200 | CGAGACTTAAATATGATACACCACTTAACCTGTGGGGGGGACCCCAATCTGGACTCCTTCCCC | 1359 |
| 1211 | CGAGACTTAAATATGATACACCACTTAACCTGTGGGGGGGACCCCAATCTGGACTCCTTCCCC | 1470 |
| 1249 | GCCTTTGGGACATCGCAGGCCCGGGAGACGCTGCCCGCAGGCTCTGGCCAGGAGAGCTCCA | 1308 |
| 1260 | GCCTTTGGGACATCGCAGGCCCGGGAGACGCTGCCCGCAGGCTCTGGCCAGGAGAGCTCCA | 1421 |
| 1309 | GGAAGGGCACTGAGCGCTCTGCGCGGAGGCTCTGGACATCCGACGCCACGAGGGAAAGT | 1368 |
| 1320 | GGAAGGGCACTGAGCGCTCTGCGCGGAGGCTCTGGACATCCGACGCCACGAGGGAAAGT | 1411 |
| 1369 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1428 |
| 1380 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1479 |
| 1418 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1530 |
| 1429 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1581 |
| 1467 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1632 |
| 1478 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1683 |
| 1516 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1734 |
| 1527 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1785 |
| 1565 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1836 |
| 1576 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1887 |
| 1614 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1938 |
| 1625 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 1989 |
| 1663 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2040 |
| 1674 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2091 |
| 1712 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2142 |
| 1723 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2193 |
| 1761 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2244 |
| 1772 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2295 |
| 1810 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2346 |
| 1821 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2397 |
| 1859 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2448 |
| 1870 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2499 |
| 1908 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2550 |
| 1919 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2601 |
| 1957 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2652 |
| 1968 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2703 |
| 2006 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2754 |
| 2017 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2805 |
| 2055 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2856 |
| 2066 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2907 |
| 2104 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 2958 |
| 2115 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3009 |
| 2153 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3060 |
| 2164 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3111 |
| 2202 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3162 |
| 2213 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3213 |
| 2251 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3264 |
| 2262 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3315 |
| 2300 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3366 |
| 2311 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3417 |
| 2349 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3468 |
| 2360 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3519 |
| 2398 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3570 |
| 2409 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3621 |
| 2447 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3672 |
| 2458 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3723 |
| 2496 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3774 |
| 2507 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3825 |
| 2545 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3876 |
| 2556 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3927 |
| 2594 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 3978 |
| 2605 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4029 |
| 2643 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4080 |
| 2654 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4131 |
| 2692 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4182 |
| 2703 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4233 |
| 2741 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4284 |
| 2752 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4335 |
| 2790 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4386 |
| 2801 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4437 |
| 2839 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4488 |
| 2850 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4539 |
| 2888 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4590 |
| 2899 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4641 |
| 2937 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4692 |
| 2948 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 4743 |
| 2986 | CTCTCGGGCGGATCTGTAATTAACCTTTTCTTTCTTTCTTTTAAATAAAAAAAAAAAAAA | 47 |

```

FEATURES
  source
POLYMER: Location/Qualifiers
1..737
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="UI-CF-EN1-seq-m-17-0-UI"
  /tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
  /dev_stage="Adult"
  /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
  /clone_lib="UI-CF-EN1"
  /notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 13:19:30 ; Search time 653 Seconds
(without alignments)
10067.131 Million cell updates/sec

Title: US-09-936-456-1

Perfect score: 1441

Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2995936 seqs, 228098010 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US06_PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------------------------|
| 1 | 1413 | 98.1 | 1421 | 15 | US-10-187-657-2 Sequence 2, Appli |
| 2 | 1404 | 97.4 | 1420 | 15 | US-10-187-657-4 Sequence 4, Appli |
| 3 | 1285 | 89.2 | 1474 | 16 | US-10-264-237-1097 Sequence 1097, Ap |
| 4 | 392 | 27.2 | 2684 | 16 | US-10-094-749-1195 Sequence 1195, Ap |
| 5 | 313 | 21.7 | 522 | 9 | US-10-187-657-5 Sequence 5, Appli |
| 6 | 187 | 13.0 | 232 | 9 | US-09-783-590-9492 Sequence 9492, Ap |
| 7 | 136 | 9.4 | 346 | 15 | US-10-187-657-3 Sequence 3, Appli |
| 8 | 119 | 8.3 | 450 | 15 | US-10-187-657-7 Sequence 7, Appli |
| 9 | 107 | 7.4 | 439 | 13 | US-10-027-632-91970 Sequence 91970, A |
| 10 | 107 | 7.4 | 439 | 16 | US-10-027-632-91970 Sequence 91970, A |
| 11 | 59 | 4.1 | 512 | 15 | US-10-106-698-3277 Sequence 3277, Ap |
| 12 | 56 | 3.9 | 506 | 15 | US-10-187-657-6 Sequence 6, Appli |
| 13 | 38 | 2.6 | 999 | 12 | US-09-876-143-1614 Sequence 1614, Ap |
| 14 | 38 | 2.6 | 1398 | 13 | US-10-424-599-64755 Sequence 64755, A |

Sequence 18, Appli
Sequence 75074, A
Sequence 842, App
Sequence 2174, Ap
Sequence 41481, A
Sequence 8115, Ap
Sequence 1, Appli
Sequence 234, App
Sequence 52594, A
Sequence 52594, A
Sequence 48927, A
Sequence 43341, A
Sequence 133773, A
Sequence 50281, A
Sequence 1341, A
Sequence 13053, A
Sequence 8647, Ap
Sequence 8179, Ap
Sequence 22997, A
Sequence 120735, A
Sequence 96207, A
Sequence 17387, A
Sequence 17007, A
Sequence 3130, Ap
Sequence 10941, A
Sequence 122157, A
Sequence 260991, A
Sequence 260991, A
Sequence 5, Appli
Sequence 91053, A

ALIGNMENTS

RESULT 1

US-10-187-657-2

; Sequence 2, Application US/10187657

; Publication No. US20030068311A1

; GENERAL INFORMATION:

; APPLICANT: Lasek, Amy K.W.

; APPLICANT: Azimzai, Yalda

; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER

; FILE REFERENCE: PV-0009 CIP

; CURRENT APPLICATION NUMBER: US/10/187,657

; CURRENT FILING DATE: 2002-07-01

; PRIOR APPLICATION NUMBER: PCT/US00/07817

; PRIOR FILING DATE: 2000-03-22

; PRIOR APPLICATION NUMBER: 60/139,565

; PRIOR FILING DATE: 1999-06-16

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PERL Program

; SEQ ID NO 2

; LENGTH: 1421

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20030068311A1 4901066CB1

US-10-187-657-2

Query Match 98.1%; Score 1413; DB 15; Length 1421;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTAACGGCTACAGACAGTGAAGTAATAGTTGCTGCGGCTAGAAAACCTCTGCG 60

DB 9 AAAGTAACGGCTACAGACAGTGAAGTAATAGTTGCTGCGGCTAGAAAACCTCTGCG 68

QY 61 GTACCAACCCAGACGGCTTGAGAGCAGCCACCTCCAGCTTCCTTAACGAGAGGTGCA 120

Db 69 GTACCAACCCAGAGCGTTGAGAGAGCCCACTCCACGCTTCCTTAACGGAGAGGTGCA 128
QY 121 GGA CT CAG A C T T C A C C A G C C C A C T C G G T C C C A G C C T T G T A C G A A A G A G A C C C C A A G G A C 180
Db 129 GGA C T C A G A C T T C A C C A G C C C A C T C G G T C C C A G C C T T G T A C G A A A G A G A C C C C A A G G A C 188
QY 181 G G C T C T C C G G T C C A G C C A G C C C C A G C T T G T G C C T T G C C T G C C G C C T C G T C A G C 240
Db 189 G G C T C T C C G G T C C A G C C A G C C C C A G C T T G T G C C T T G C C T G C C G C C T C G T G C A G C 248
QY 241 A C T C G C C G C G C G T A G C A C C T G T G G A A C G G C T A C T G C C T T T T T A C C C C C A G C C C 300
Db 249 A C T C G C C G C G C G T A G C A C C T G T G G A A C G G C T A C T G C C T T T T T A C C C C C A G C C C 308
QY 301 C G C A C T C C G C A G C C T C A G C G T C C A C T G C T C A T G T T A T T C T A G T T T T T G C T C T A 360
Db 309 C G C A T T G C C G C A G C C T T C A G C G T T C A C T G C T C A T G T T A T T C T A G T T T T T G C T C T A 368
QY 361 G C A G C A A G C T T C C T G C T C A T T C T T G C C G G G A T C C G T G G C A C T C G C G C T G T T T G G T T G 420
Db 369 G C A G C A A G C T T C C T G C T C A T T C T T G C C G G G A T C C G T G G C A C T C G C G C T G T T T G G T T G 428
QY 421 G T C A G A G T T C T T C A G T C T G T T C A T A G G C G A G A A T T G T G C T T G C A C T T C A G T G C A 480
Db 429 G T C A G A G T T C T T C A G T C T G T T C A T A G G C G A G A A T T G T G C T T G C A C T T C A G T G C A 488
QY 481 G A A T G T T C G T G G T A C A G A C A C A C A C A C T C T A A A A C C T T C A G C G A G C G C C 540
Db 489 G A A T G T T C G T G G T A C A G A C A C A C A C A C T C T A A A A C C T T C A G C G A G C G C C 548
QY 541 G T T A C A G C C G T C G C T G T G T G G C C T G A G G C G A T T A T T A T T A C A C T C A C A G G 600
Db 549 G T T A C A G C C G T C G C T G T G T G G C C T G A G G C A T T A A T A T T A C A C T C A C A G G 608
QY 601 A C C C C A G T C A T C A G T G A A G A C C A T T G A C T A A A C A G C A G T T C A C C T G G C C T G 660
Db 609 A C C C C A G T C A T C A G T G A A G A C C A T T G A C T A A A C A G C A G T T C A C C T G G C C T G 668
QY 661 A A A G A A T T A C C C G G A G T A G C G A A C G C A C T G G A A G G G T G C C G A C C C A G T G 720
Db 669 A A A G A A T T A C C C G G A G T A G C G A A C G C A C T G G A A G G G T G C C G A C C C A G T G 728
QY 721 C T C A C T C G C G A G A G T T C A C C A G A G T A G C C T T G G G C C T G T A C C A C A G A T T A C C A C 780
Db 729 C T C A C T C G C G A G A G T T C A C C A G A G T A G C C T T G G G C C T G T A C C A C A G A T T A C C A C 788
QY 781 C T C G C G A C A C T A C G C T C G G C A C G C T A T A G G T G C G T T C T T C T T G C T C T C C 840
Db 789 C T C G C G A C A C T A C G C T C G G C A C G C T A T A G G T G C G T T C T T C T T G C T C T C C 848
QY 841 A A C G T G C T C T C C A C C G C C C C C C C T C T A C G A G G C C T G C A C T G C T G A C C A C G G A 900
Db 849 A A C G T G C T C T C C A C C G C C C C C C C C T C T A C G A G G C C T G C A C T G C T G A C C A C G G A 908
QY 901 G C T T C G C G C T T T C G G G G T C T T G C C T T G C C C T C C A T C T A G C G T G C C G C T C T G C C G 960
Db 909 G C C T T C G C G C T T T C G G G G T C T T G C C T T G C C C T C C A T C T C T A G C G T G C C G C T C G C C G 968
QY 961 C T C G C C T A G G C T C T C G C G C T C A C C A C A C A G C G G C C C C C T T C T G G G T C A C G T G 1020
Db 969 C T C G C C T A G G C T C T C G C G C T C A C C A C A C A G C G G C C C C C T T C T G G G T C A C G C T G 1028
QY 1021 G C A A C C G G C C T C T G T G C C T C T T C C T C G A G G G C C G T G T A G T C T C A G T A T T T C G 1080
Db 1029 G C A A C C G G C C T C T G T G C C T C T T C C T C G A G G G C C G T G T A G T C T C A G T A T T T C G 1088
QY 1081 C C A G C G C T C T T C C A C C C T T T G A A C A A A G C C A A G C A C T G C A G C C A G G A G A G G G 1140
Db 1089 C C A G C G C T C T T C G A C C C T T T G A C C A A A G C C A A G C A C T G C A G C C A G G A G A G G G 1148
QY 1141 G G C T C A C T T A T C T C G G G A C C C A C T C A C A G C A G C C C C T C T C C A G A C T T A A A 1200
Db 1149 G G C T C A C C T T A T C T C G G G A C C C A C T G A A A G C A G C C C C C T C T C C A G A C T T A A A 1208

RESULT 2

US-10-187-657-4

; Sequence 4, Application US/10187657

; Publication No. US20030068311A1

; GENERAL INFORMATION:

; APPLICANT: Iasek, Amy K.W.

; APPLICANT: Baughn, Mariah R.

; APPLICANT: Azimzai, Yalda

; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER

; FILE REFERENCE: PV-0009 CIP

; CURRENT APPLICATION NUMBER: US/10/187,657

; PRIOR FILING DATE: 2002-07-01

; PRIOR APPLICATION NUMBER: PCT/US00/07817

; PRIOR FILING DATE: 2000-03-22

; PRIOR APPLICATION NUMBER: 60/139,565

; PRIOR FILING DATE: 1999-06-16

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PERL Program

; SEQ ID NO 4

; LENGTH: 1420

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20030068311A1 3221661CA2

US-10-187-657-4

Query Match 97.4%; Score 1404; DB 15; Length 1420;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGCTAGAAAAAATCTGTGG 60

Db 9 AAAGTAAACGGCTACAGACAGTGAAGAAATAGTTTCGCTCGCGGCTAGAAAAAATCTGTGG 68

QY 61 GTACCAACCCAGAGCGTTGAGAGAGCCCACTCCACGCTTCCTTAACGGAGAGGTGCA 120

Db 69 GTACCAACCCAGAGCGTTGAGAGAGCCCACTCCACGCTTCCTTAACGGAGAGGTGCA 128

QY 121 GGA CT CAG A C T T C A C C A G C C C A C T C G G T C C C A G C C T T G T A C G A A A G A G A C C C C A A G G A C 180

Db 129 GGA CT CAG A C T T C A C C A G C C C A C T C G G T C C C A G C C T T G T A C G A A A G A G A C C C C A A G G A C 188

QY 181 G G C T C T C C G G T C C A G C C A G C C C C A G C T T G T G C C T T G C C C C G C C T C G T G C A G C 240

Db 189 G G C T C T C C G G T C C A G C C A G C C C C A G C T T G T G C C T T G C C C C G C C T C G T G C A G C 248

QY 241 A C T C G C C G C G C G T C A G C A T G A C C C T G T G A A C G G C T A C T G C C T T T T T A C C C C C A G C C C 300

Db 249 A C T C G C C G C G C G T C A G C A T G A C C C T G T G A A C G G C T A C T G C C T T T T T A C C C C C A G C C C 308

QY 301 C G C A C T C C G C A G C C T T C A G C G T T C C A C T G C T C A T G C T A T T C T A G T T T T T G C T C T A 360

Db 309 C G C A T C C G C A G C C T T C A G C G T T C C A C T G C T C A T G C T A T T C T A G T T T T T G C T C T A 368

361 GCAGCAGCTTCTGCTCATCTTCCGGGGATCCGTGGCCACTCGCGCTGTTGGTTG 420
Db GCAGCAGCTTCTGCTCATCTTCCGGGGATCCGTGGCCACTCGCGCTGTTGGTTG 428
QY GTGAGAGTTCTTCTCAGTCTGTTTCATAGCGCAGAAATTTGGTGTGCTCACTTCAGTGCA 480
Db GTGAGAGTTCTTCTCAGTCTGTTTCATAGCGCAGAAATTTGGTGTGCTCACTTCAGTGCA 488
QY GAATGGTTCTGGGTACAGTGAACCAACACATCTTCAAGGCTTCAGCGCGCGCGC 540
Db GAATGGTTCTGGGTACAGTGAACCAACACATCTTCAAGGCTTCAGCGCGCGCGC 548
QY GTTACAGCCCGTGTCCGTCTGCTCGTGGCGCTGAGGCGCATTAATATTACACTCACAGG 600
Db GTTACAGCCCGTGTCCGTCTGCTCGTGGCGCTGAGGCGCATTAATATTACACTCACAGG 608
QY ACCCAGTGCATCACTGAAAGAGACCATTTGACTACAAAGAGAGTTCACTTCGGGCTG 660
Db ACCCAGTGCATCACTGAAAGAGACCATTTGACTACAAAGAGAGTTCACTTCGGGCTG 668
QY AAAGAGAAATTACGCGCGAGTACGCAACGCACTGAGAGAGGGCTCGCGGACCCAGTG 720
Db AAAGAGAAATTACGCGCGAGTACGCAACGCACTGAGAGAGGGCTCGCGGACCCAGTG 728
QY CTCTACCTGGCGGAGAGTTCAACCGAGTAGCCCTTCCGCTGTATACCAACAGTACAC 780
Db CTCTACCTGGCGGAGAGTTCAACCGAGTAGCCCTTCCGCTGTATACCAACAGTACAC 788
QY CTGGCGGACACTACGCTCGGCGCAAGCTATGGGTGGGTCTGCTTCTGGCTCTCTCC 840
Db CTGGCGGACACTACGCTCGGCGCAAGCTATGGGTGGGTCTGCTTCTGGCTCTCTCC 848
QY AACGTGCTGCTCTCCACGCGCGCCGCTCTACGAGAGCCCTGGCACTCTGACCAACCGGA 900
Db AACGTGCTGCTCTCCACGCGCGCCGCTCTACGAGAGCCCTGGCACTCTGACCAACCGGA 908
QY GCCTTCGGGCTCTTCGGGCTTTCGCTTGGCTTCCATCTCTAGGCTGCGCTTCGCGC 960
Db GCCTTCGGGCTCTTCGGGCTTTCGCTTGGCTTCCATCTCTAGGCTGCGCTTCGCGC 968
QY CTCGCCCTTAGGCTCTTCGCGCTCAACCACTAGTACGCGCGCCCTCTGAGTCAAGCTG 1020
Db CTCGCCCTTAGGCTCTTCGCGCTCAACCACTAGTACGCGCGCCCTCTGAGTCAAGCTG 1028
QY GCAACCGGCTCTTCGCGCTCTTCGAGGGGCGCTGGTGGTCTCAGTATGTTGG 1080
Db GCAACCGGCTCTTCGCGCTCTTCGAGGGGCGCTGGTGGTCTCAGTATGTTGG 1088
QY CCCAGGCTCTTCGCGCTCTTCGAGGCGCGCTGGTGGTCTCAGTATGTTGG 1140
Db CCCAGGCTCTTCGCGCTCTTCGAGGCGCGCTGGTGGTCTCAGTATGTTGG 1148
QY GGCTCACTCTTATCTCGGACCCACTGCAACAGCAGGCGGCTCTCCAGACTTAAAA 1200
Db GGCTCACTCTTATCTCGGACCCACTGCAACAGCAGGCGGCTCTCCAGACTTAAAA 1208
QY TGTATCACTTAACTGTAGGGGACCCCAATCTGGACTCTTCCGCGCTTGGGACAT 1260
Db TGTATCACTTAACTGTAGGGGACCCCAATCTGGACTCTTCCGCGCTTGGGACAT 1268
QY CCGAGCGCGGAGCAGTGCCTCGCGCCAGGCTTGGGCGCAGAGAGCTTCCAGGAGGCACTG 1320
Db CCGAGCGCGGAGCAGTGCCTCGCGCCAGGCTTGGGCGCAGAGAGCTTCCAGGAGGCACTG 1328
QY ACCGCTGTGGCGGAGGCTCGGACATCCGAGGACCCAGGGAAGTCTCTCTGGGGGA 1380
Db ACCGCTGTGGCGGAGGCTCGGACATCCGAGGACCCAGGGAAGTCTCTCTGGGGGA 1388
QY TCTGTAAATAAACCTTTTTTTCTT 1404
Db TCTGTAAATAAACCTTTTTTTCTT 1412

RESULT 3

US-10-264-237-1097
; Sequence 1097, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Biese et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1097
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-237-1097

Query Match 89.2%; Score 1285; DB 16; Length 1474;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1435; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AAAGTAAACGGCTACAGACAGTGAATAAGTTTGGCTCGCGGCTAGAAAACTCTGTG 60
Db 27 AAAGTAAACGGCTACAGACAGTGAATAAGTTTGGCTCGCGGCTAGAAAACTCTGTG 86
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAAACGAGAGGTGCA 120
Db 87 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCAGCTTCTTAAACGAGAGGTGCA 146
QY 121 GGACTCAGACTTTCACACAGCCCACTCGGTCCAGCTTGTACGCAAGAGAGAGCCAGAC 180
Db 147 GEAATCAGACTTTCACACAGCCCACTCGGTCCAGCTTGTACGCAAGAGAGAGCCAGAC 206
QY 181 GCGCTCTCCCGCTCAGAGCAGCCCGAGCTTGTGCTTGCCTCCCGCTCGCTGAGC 240
Db 207 GCGCTCTCCCGCTCAGAGCAGCCCGAGCTTGTGCTTGCCTCCCGCTCGCTGAGC 266
QY 241 ACTCGCGCGGTGTCAGCATGACCTGTGAAACGCGCTACTGCTCTTTTACCCCGAGCC 300
Db 267 ACTCGCGCGGTGTCAGCATGACCTGTGAAACGCGCTACTGCTCTTTTACCCCGAGCC 326
QY 301 CGCATCCCGAGCTTTCAGCGTTCCACTGCTCATGCTTATCTAGTGTGTTTGGCTCTA 360
Db 327 CGCATCCCGAGCTTTCAGCGTTCCACTGCTCATGCTTATCTAGTGTGTTTGGCTCTA 386
QY 361 GCAGAGCTTCTGCTCATCTTCCGCGGATCCGTGCGCACTCGCTGTTTGGTTG 420
Db 387 GCAGAGCTTCTGCTCATCTTCCGCGGATCCGTGCGCACTCGCTGTTTGGTTG 446
QY 421 GTGAGAGTTCTTCTCAGTCTGTTTCATAGCGCAGAAATTTGGTGTGCTTCACTCAGTGCA 480
Db 447 GTGAGAGTTCTTCTCAGTCTGTTTCATAGCGCAGAAATTTGGTGTGCTTCACTCAGTGCA 506
QY 481 GAATGGTTCTGGGTACAGTGAACCAACATCTTCAAGGCTTCAGCGCGCGCGC 540
Db 507 GAATGGTTCTGGGTACAGTGAACCAACATCTTCAAGGCTTCAGCGCGCGCGC 566
QY 541 GTTACAGCCCGTGTCCGTCTGCTCGTGGGCTCGAGGCGCATTAATATTACACTCACAGG 600
Db 567 GTTACAGCCCGTGTCCGTCTGCTCGTGGGCTCGAGGCGCATTAATATTACACTCACAGG 626
QY 601 ACCCAGTGCATCACTGAAAGAGACCATTTGACTACAAAGAGAGTTCACTTCGGGCTG 660
Db 627 ACCCAGTGCATCACTGAAAGAGACCATTTGACTACAAAGAGAGTTCACTTCGGGCTG 686
QY 661 AAAGAGAAATTACGCGCGAGTACGCAACGCACTGAGAGAGGGCTCGCGGACCCAGTG 720
Db 687 AAAGAGAAATTACGCGCGAGTACGCAACGCACTGAGAGAGGGCTCGCGGACCCAGTG 746


```

; OTHER INFORMATION: Incyte ID No. US20030068311A1 322166186
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 455, 480, 483
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-5

Query Match
Best Local Similarity 21.7%; Score 313; DB 15; Length 522;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGTAACGGGTACAGACAGTGAAGAAATAGTTTCGCTCGCCGGGTAGAAAACTCTGTGCG 60
Db 9 AAAGTAACGGGTACAGACAGTGAAGAAATAGTTTCGCTCGCCGGGTAGAAAACTCTGTGCG 68

Qy 61 GTACCAACCCAGAGCGTTGAGACAGCCACCTCCACGCTTCCTTAACGAGAGGTGCA 120
Db 69 GTACCAACCCAGAGCGTTGAGACAGCCACCTCCACGCTTCCTTAACGAGAGGTGCA 128

Qy 121 GGACTCAGACTTCACGACCCACCTCGTCCAGCCTTTGTACGCAAGAGAGGCCAAGGAC 180
Db 129 GGACTCAGACTTCACGACCCACCTCGTCCAGCCTTTGTACGCAAGAGAGGCCAAGGAC 188

Qy 181 GCGCTCTCCCGGTCCAGGAGCCCGACCTTGCTGGCTTGGCTGCGCGCTGGTGCAGC 240
Db 189 GCGCTCTCCCGGTCCAGGAGCCCGACCTTGCTGGCTTGGCTGCGCGCTGGTGCAGC 248

Qy 241 ACTCGCGCGGTGAGCATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGACGCC 300
Db 249 ACTCGCGCGGTGAGCATGACCTGTGGAACGGCGTACTGCTTTTATACCCCGACGCC 308

Qy 301 CGGCATGCCGAG 313
Db 309 CGGCATGCCGAG 321

RESULT 6
US-09-783-590-9492
; Sequence 9492, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9492
; LENGTH: 232
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (202)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (206)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (229)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9492

Query Match
Best Local Similarity 13.0%; Score 187; DB 9; Length 232;

```

```

Best Local Similarity 100.0%; Pred. No. 2.7e-83;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAAGTAACGGGTACAGACAGTGAAGAAATAGTTTCGCTCGCCGGGTAGAAAACTCTGTGCG 60
Db 15 AAAGTAACGGGTACAGACAGTGAAGAAATAGTTTCGCTCGCCGGGTAGAAAACTCTGTGCG 74

Qy 61 GTACCAACCCAGAGCGTTGAGACAGCCACCTCCACGCTTCCTTAACGAGAGGTGCA 120
Db 75 GTACCAACCCAGAGCGTTGAGACAGCCACCTCCACGCTTCCTTAACGAGAGGTGCA 134

Qy 121 GGACTCAGACTTCACGACCCACCTCGTCCAGCCTTTGTACGCAAGAGAGGCCAAGGAC 180
Db 135 GGACTCAGACTTCACGACCCACCTCGTCCAGCCTTTGTACGCAAGAGAGGCCAAGGAC 194

Qy 181 GCGCTCT 187
Db 195 GCGCTCT 201

RESULT 7
US-10-187-657-3
; Sequence 3, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Lasek, Amy K.W.
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 346
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 1752794F6
; NAME/KEY: unsure
; LOCATION: 91, 188, 206, 291, 337
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-3

Query Match
Best Local Similarity 9.4%; Score 136; DB 15; Length 346;
Matches 186; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 927 CTTGGGCTCCATCTCTAGCGTGCCTCTGCGCGCTCGCTAGGCTCTCTCGCGCTCAC 986
Db 1 CTTGGGCTCCATCTCTAGCGTGCCTCTGCGCGCTCGCTAGGCTCTCTCGCGCTCAC 60

Qy 987 CACTCAGTACGGCGCGCTTCTGGGTCACTGGCAACCGGCTCTGCTCTTCTCT 1046
Db 61 CACTCAGTACGGCGCGCTTCTGGGTCACTGGCAACCGGCTCTGCTCTTCTCT 120

Qy 1047 CGGAGGGCGGTGCTGAGTCTCCAGTATGTTCCGCCAGCGCTCTTCGACCCCTTCTGGA 1106
Db 121 CGGAGGGCGGTGCTGAGTCTCCAGTATGTTCCGCCAGCGCTCTTCGACCCCTTCTGGA 180

Qy 1107 CCAAGC 1113
Db 181 CCAAGC 187

RESULT 8

```

```
US-10-187-657-7
; Sequence 7, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzai, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 SEQID4510D1
;
; NAME/KEY: unsure
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-7

Query Match      8.3%; Score 119; DB 15; Length 450;
Best Local Similarity 100.0%; Pred. No. 3.6e-49;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 CTGACGAGACCATGACTACAGCAGCTTCACTGGGCTCTGAAAGAGAAATTACGCC 675
DB 21 CTGACGAGACCATGACTACAGCAGCTTCACTGGGCTCTGAAAGAGAAATTACGCC 80
QY 676 GCGGAGTACGCGAACGACCTGGAGAGGGGCTGCGGACCGACGTCCTACCTGGCGGA 734
DB 81 GCGGAGTACGCGAACGACCTGGAGAGGGGCTGCGGACCGACGTCCTACCTGGCGGA 139

RESULT 9
US-10-027-632-91970
; Sequence 91970, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/156,358
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91970
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G

US-10-027-632-91970

Query Match      7.4%; Score 107; DB 16; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.9e-43;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 892 ACCACCGAGCCTTCGCGCTCTTCGCGGCTTCGCGCTTCATCTTAGCGTGGCG 951
DB 214 ACCACCGAGCCTTCGCGCTCTTCGCGGCTTCGCGCTTCATCTTAGCGTGGCG 273
QY 952 CTCTGCCCGCTCCGCTAGGCTCTCCGCGCTCACCACCTCAGTACGG 998
DB 274 CTCTGCCCGCTCCGCTAGGCTCTCCGCGCTCACCACCTCAGTACGG 320

RESULT 10
US-10-027-632-91970
; Sequence 91970, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91970
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G

US-10-027-632-91970

Query Match      7.4%; Score 107; DB 16; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.9e-43;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 892 ACCACCGAGCCTTCGCGCTCTTCGCGGCTTCGCGCTTCATCTTAGCGTGGCG 951
DB 214 ACCACCGAGCCTTCGCGCTCTTCGCGGCTTCGCGCTTCATCTTAGCGTGGCG 273
QY 952 CTCTGCCCGCTCCGCTAGGCTCTCCGCGCTCACCACCTCAGTACGG 998
DB 274 CTCTGCCCGCTCCGCTAGGCTCTCCGCGCTCACCACCTCAGTACGG 320

RESULT 11
US-10-106-698-3277
; Sequence 3277, Application US/10106698
```

```
US-10-027-632-91970
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-91970

Query Match      7.4%; Score 107; DB 13; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.9e-43;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 892 ACCACCGAGCCTTCGCGCTCTTCGCGGCTTCGCGCTTCATCTTAGCGTGGCG 951
DB 214 ACCACCGAGCCTTCGCGCTCTTCGCGGCTTCGCGCTTCATCTTAGCGTGGCG 273
QY 952 CTCTGCCCGCTCCGCTAGGCTCTCCGCGCTCACCACCTCAGTACGG 998
DB 274 CTCTGCCCGCTCCGCTAGGCTCTCCGCGCTCACCACCTCAGTACGG 320

RESULT 10
US-10-027-632-91970
; Sequence 91970, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91970
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G

US-10-027-632-91970

Query Match      7.4%; Score 107; DB 16; Length 439;
Best Local Similarity 100.0%; Pred. No. 3.9e-43;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 892 ACCACCGAGCCTTCGCGCTCTTCGCGGCTTCGCGCTTCATCTTAGCGTGGCG 951
DB 214 ACCACCGAGCCTTCGCGCTCTTCGCGGCTTCGCGCTTCATCTTAGCGTGGCG 273
QY 952 CTCTGCCCGCTCCGCTAGGCTCTCCGCGCTCACCACCTCAGTACGG 998
DB 274 CTCTGCCCGCTCCGCTAGGCTCTCCGCGCTCACCACCTCAGTACGG 320

RESULT 11
US-10-106-698-3277
; Sequence 3277, Application US/10106698
```

```

; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA003P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 3277
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (392)..(392)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (453)..(453)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (461)..(461)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (493)..(493)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (499)..(499)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (509)..(509)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-106-698-3277
;
; Query Match
; Best Local Similarity 100.0%; Pred. No. 4.9e-19; Length 512;
; Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 405 GCGCTGGTTTGGTGGTGAGATCTCTCAGTCTGTTCTAGCGCGAGAAATTTGGG 463
; DB 220 GCGCTGGTTTGGTGGTGAGATCTCTCAGTCTGTTCTAGCGCGAGAAATTTGGG 278
;
; RESULT 12
; US-10-187-657-6
; Sequence 6, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Azimzal, Valda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CIP
; CURRENT APPLICATION NUMBER: US/10/187,657
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 506
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```

; OTHER INFORMATION: Incyte ID No. US20030068311A1 SBQA03652D1
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 87, 89, 93, 95, 124, 178, 227, 253, 262, 279, 285, 293-294, 298,
; LOCATION: 300, 303, 312, 319-321, 323, 325, 327, 332-333, 336, 343-344, 349, 377, 381
; LOCATION: 405, 409-410, 412, 415, 420, 423, 429, 432, 436-437, 444-445, 448, 452, 454
; LOCATION: 466, 469-471, 479, 483, 485, 492, 494, 499, 501, 504
; OTHER INFORMATION: a, t, c, g, or other
; US-10-187-657-6
;
; Query Match
; Best Local Similarity 100.0%; Pred. No. 1.6e-17; Length 506;
; Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 479 CAGATGGTTGGTGGTACAGTGAACACCAACACATCTCTACAAAGCCTTCAGCGCA 534
; DB 31 CAGATGGTTGGTGGTACAGTGAACACCAACACATCTCTACAAAGCCTTCAGCGCA 86
;
; RESULT 13
; US-09-876-143-1614
; Sequence 1614, Application US/09876143
; Publication No. US20040081958A1
; GENERAL INFORMATION:
; APPLICANT: Infogen Inc.
; APPLICANT: EILERTSEN, KENNETH J.
; APPLICANT: FRIESTER-GENSKOW, MARTHA
; APPLICANT: CHILDS, LYNETTE
; APPLICANT: FORSYTHE, TODD
; APPLICANT: BISHOP, MICHAEL D.
; TITLE OF INVENTION: IDENTIFICATION AND USE OF MOLECULAR MARKERS INDICATING
; TITLE OF INVENTION: CELLULAR REPROGRAMMING
; FILE REFERENCE: 028040-0202
; CURRENT APPLICATION NUMBER: US/09/876,143
; CURRENT FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/209,874
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 1744
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1614
; LENGTH: 999
; TYPE: DNA
; ORGANISM: Bovine
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(999)
; OTHER INFORMATION: n is a, c, g, or t
; US-09-876-143-1614
;
; Query Match
; Best Local Similarity 100.0%; Pred. No. 1.6e-08; Length 999;
; Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1404 TTGTTTTTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 1441
; DB 729 TTGTTTTTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 766
;
; RESULT 14
; US-10-424-599-64755
; Sequence 64755, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

```

Fri Jun 4 16:04:06 2004

```

; SEQ ID NO 64755
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29487C.1
US-10-424-599-64755

Query Match      2.6%; Score 38; DB 13; Length 1398;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1404 TTGCTTTTAAAAA 1441
Db 1325 TTGCTTTTAAAAA 1362

```

```

RESULT 15
US-10-116-255-18
; Sequence 18, Application US/10116255
; Publication No. US20030036646A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: Uncoupling Protein Polynucleotides, Polypeptides, and
; FILE REFERENCE: PTO09Pl Antibodies
; CURRENT APPLICATION NUMBER: US/10/116,255
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 09/685,897
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: PCT/US00/09534
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/128,701
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: 60/142,821
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/149,448
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/164,751
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 2165
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-255-18

```

```

Query Match      2.6%; Score 38; DB 15; Length 2165;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1404 TTGCTTTTAAAAA 1441
Db 2128 TTGCTTTTAAAAA 2165

```

Search completed: June 3, 2004, 16:28:19
Job time : 639 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 08:19:31 ; Search time 637 Seconds
(without alignments)
9610.131 Million cell updates/sec

Title: US-09-936-456-1
Perfect score: 1441

Sequence: 1 aaagtaacggctacagacag.....aaaaaaaaaaaaaaaaaaaaa 1441

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_29Jan04.*
- 1: Geneseqn1980s.*
 - 2: Geneseqn1990s.*
 - 3: Geneseqn2000s.*
 - 4: Geneseqn2001as.*
 - 5: Geneseqn2001bs.*
 - 6: Geneseqn2002s.*
 - 7: Geneseqn2003as.*
 - 8: Geneseqn2003bs.*
 - 9: Geneseqn2003cs.*
 - 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 1441 | 100.0 | 1441 | 3 | AAA94623 Human CAS |
| 2 | 1436 | 99.7 | 1491 | 9 | ADD19228 Human cDN |
| 3 | 1435.6 | 99.6 | 1474 | 6 | AB190535 Human pol |
| 4 | 1413 | 98.1 | 1421 | 3 | AAA96505 cDNA enco |
| 5 | 1392.8 | 96.7 | 1460 | 3 | AAC95564 Human sec |
| 6 | 1392.8 | 96.7 | 1460 | 3 | ABZ67235 Human sec |
| 7 | 1392.8 | 96.7 | 1460 | 7 | ABZ73640 Secreted |
| 8 | 1392.8 | 96.7 | 1460 | 9 | ADC20289 Human sec |
| 9 | 465.4 | 32.3 | 498 | 3 | AAA94624 Human CAS |
| 10 | 405.4 | 28.1 | 5033 | 3 | ABZ68115 Human sec |
| 11 | 405.4 | 28.1 | 5033 | 7 | ABZ74587 Secreted |
| 12 | 405.4 | 28.1 | 5033 | 9 | ADC21005 Human sec |
| 13 | 405 | 28.1 | 580 | 9 | ADD19187 Human cDN |
| 14 | 401.8 | 27.9 | 406 | 5 | AAS65653 DNA enco |
| 15 | 396 | 27.5 | 2684 | 7 | ADA53627 Human cod |
| 16 | 356.4 | 24.7 | 1636 | 6 | ABN59923 Novel hum |
| 17 | 354.8 | 24.6 | 1029 | 5 | AAP45134 Human INT |
| 18 | 354.8 | 24.6 | 1029 | 7 | ACD66787 Secreted |
| 19 | 354.8 | 24.6 | 2133 | 5 | AAP45133 Human INT |
| 20 | 354.8 | 24.6 | 2133 | 7 | ABX94107 cDNA enco |
| 21 | 354.8 | 24.6 | 2133 | 7 | ACD66786 Secreted |
| 22 | 245.8 | 17.1 | 2684 | 7 | ADA53627 Human cod |
| 23 | 161.8 | 11.2 | 1338 | 5 | AAS67707 DNA enco |

| | | | | | | |
|----|------|-----|-------|---|----------|--------------------|
| 24 | 65.6 | 4.6 | 1602 | 4 | ABL21213 | ABL21213 Drosophil |
| 25 | 65.6 | 4.6 | 1808 | 4 | ABL07775 | ABL07775 Drosophil |
| 26 | 64 | 4.4 | 512 | 4 | AAB36185 | AAB36185 Human col |
| 27 | 60.6 | 4.2 | 2000 | 7 | ADA71938 | ADA71938 Rice gene |
| 28 | 52.6 | 3.7 | 134 | 7 | ABZ68116 | ABZ68116 Human sec |
| 29 | 52.6 | 3.7 | 134 | 7 | ABZ74588 | ABZ74588 Secreted |
| 30 | 52.6 | 3.7 | 134 | 9 | ADC21006 | ADC21006 Human sec |
| 31 | 52 | 3.6 | 2000 | 7 | ADA71938 | ADA71938 Rice gene |
| 32 | 49 | 3.4 | 496 | 5 | ABV56664 | ABV56664 Human pro |
| 33 | 48.6 | 3.4 | 1400 | 4 | AAS01017 | AAS01017 Maize dis |
| 34 | 47.8 | 3.3 | 396 | 3 | AAB34971 | AAB34971 Wheat bet |
| 35 | 47.6 | 3.3 | 277 | 7 | ABX47508 | ABX47508 Bovine ES |
| 36 | 47.4 | 3.3 | 915 | 3 | AAC60068 | AAC60068 Human sec |
| 37 | 46.8 | 3.2 | 656 | 4 | AAL23653 | AAL23653 Human bre |
| 38 | 46.8 | 3.2 | 2484 | 5 | AAS78178 | AAS78178 DNA enco |
| 39 | 46.8 | 3.2 | 2589 | 7 | ABZ75977 | ABZ75977 Human G p |
| 40 | 46.8 | 3.2 | 14055 | 7 | AAL61170 | AAL61170 Actinosyn |
| 41 | 46.8 | 3.2 | 82746 | 7 | AAL61224 | AAL61224 Actinosyn |
| 42 | 46.4 | 3.2 | 2206 | 4 | ABK43595 | ABK43595 DNA enco |
| 43 | 46.2 | 3.2 | 441 | 3 | AAC69600 | AAC69600 Human sec |
| 44 | 46.2 | 3.2 | 531 | 8 | ADA49299 | ADA49299 Maize gen |
| 45 | 46.2 | 3.2 | 609 | 8 | ADA49313 | ADA49313 Maize gen |

ALIGNMENTS

RESULT 1
AAA94623
ID AAA94623 standard; DNA; 1441 BP.
XX
AC AAA94623;
XX
DT 11-JAN-2001 (first entry)
XX
DE Human CASB618 coding sequence.
XX
KW Epitope; human; CASB618; chromosome 15; tumour; vaccine; cancer; ovarian;
KW Colon; autoimmune disease; HLA_A0201; ss.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT CDS 259..1221
FT /*tag= a
FT /product= "Human CASB618"
XX
PN WO200053748-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-EP002048.
XX
PR 11-MAR-1999; 99GB-00005607.
PR 01-SEP-1999; 99GB-00020590.
XX
PA (SWIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Bruck CEM, Cassart J, Coche T, Vinals Y De BassolsC;
XX
DR WPI; 2000-572268/53.
XX
DR P-PSDB; AAB26325.
XX
PT New human CASB618 polypeptide, useful as a vaccine for prophylactic and
PT therapeutic treatment of cancers, particularly ovarian or colon cancer,
PT autoimmune diseases and related conditions.
XX
PS Claim 13; Page 61; 76pp; English.
XX
CC The present sequence is the coding sequence of human CASB618 protein. The
CC gene for human CASB618 is thought to be located on chromosome 15. The
CC protein encoded by the present sequence and epitopes of the CASB618
CC protein (see AAB26327 to AAB26399) are useful in diagnosing the occurrence

| | |
|---|--|
| CC of tumour cells and in vaccines for prophylactic and therapeutic | |
| CC treatment of cancers, particularly ovarian or colon cancer, autoimmune | |
| XX diseases and related conditions | |
| SQ Sequence 1441 BP; 289 A; 466 C; 392 G; 294 T; 0 U; 0 Other; | |
| Query Match | |
| Best Local Similarity 100.0%; Score 1441; DB 3; Length 1441; | |
| Matches 1441; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| Qy 1 | AAAGTAACCGCTACAGACAGTGAAGAAATAGTTTCGCTCGCCGGCTAGAAAACTCTGTGC 60 |
| Db 1 | AAAGTAACCGCTACAGACAGTGAAGAAATAGTTTCGCTCGCCGGCTAGAAAACTCTGTGC 60 |
| Qy 61 | GTACCAACCCAGAGGTTGAGACAGCCACCTTCAGCCCTTCTTACCGAGAGTGCA 120 |
| Db 61 | GTACCAACCCAGAGGTTGAGACAGCCACCTTCAGCCCTTCTTACCGAGAGTGCA 120 |
| Qy 121 | GGACTCAGACTTTCACAGCCCACTCGGTCCAGCCCTTGTACCAAGAGACGCCAAGGAC 180 |
| Db 121 | GGACTCAGACTTTCACAGCCCACTCGGTCCAGCCCTTGTACCAAGAGACGCCAAGGAC 180 |
| Qy 181 | GGCTCTCCCGGTCAGCAGCCAGCTTGTGTGGCTTGGCTTGCCTGCGCGCTCGTGCAGC 240 |
| Db 181 | GGCTCTCCCGGTCAGCAGCCAGCTTGTGTGGCTTGGCTTGCCTGCGCGCTCGTGCAGC 240 |
| Qy 241 | ACTCGGCCGGCTGACAGTACCTGTGGAAGCGTACTGCTTTTACCCCGAGGCC 300 |
| Db 241 | ACTCGGCCGGCTGACAGTACCTGTGGAAGCGTACTGCTTTTACCCCGAGGCC 300 |
| Qy 301 | CGGCATGCGCGAGGCTTCAGCGTTCACCTGCTCATGCTATCTAGTGTGTTTGTCTTA 360 |
| Db 301 | CGGCATGCGCGAGGCTTCAGCGTTCACCTGCTCATGCTATCTAGTGTGTTTGTCTTA 360 |
| Qy 361 | GCAGCAAGCTTCTGCTCATCTGCGGGATCCGTGGCCACTCGCGCTGTTTGTGTTG 420 |
| Db 361 | GCAGCAAGCTTCTGCTCATCTGCGGGATCCGTGGCCACTCGCGCTGTTTGTGTTG 420 |
| Qy 421 | GTGAGAGTTCTTCTCAGTCTGTTTCATAGCGCGAGAAATGTGGCTGTGCACTTCAGTGCA 480 |
| Db 421 | GTGAGAGTTCTTCTCAGTCTGTTTCATAGCGCGAGAAATGTGGCTGTGCACTTCAGTGCA 480 |
| Qy 481 | GAATGTTCTGTTGTTACAGTACACCAACACATCTACAAAGCCTTCAGCGCAGCGGC 540 |
| Db 481 | GAATGTTCTGTTGTTACAGTACACCAACACATCTACAAAGCCTTCAGCGCAGCGGC 540 |
| Qy 541 | GTTACAGCCGCTGCTCGTCTGCTGCGGCTGAGGCGCATTAATATTACACTCACAGG 600 |
| Db 541 | GTTACAGCCGCTGCTCGTCTGCTGCGGCTGAGGCGCATTAATATTACACTCACAGG 600 |
| Qy 601 | ACCCAGTGCATCAGTGAAGAGACATTTGACTACACAGCAGTTTACCTGGGCTCTG 660 |
| Db 601 | ACCCAGTGCATCAGTGAAGAGACATTTGACTACACAGCAGTTTACCTGGGCTCTG 660 |
| Qy 661 | AAAGAGATTACGCGCGAGTACGGAACGACCTGGAAGGGCTGCCGACCCAGTG 720 |
| Db 661 | AAAGAGATTACGCGCGAGTACGGAACGACCTGGAAGGGCTGCCGACCCAGTG 720 |
| Qy 721 | CTCTACTCGCGAGAGTTTCACCGAGTAGCCCTTGCCTGAGTACACAGTACAC 780 |
| Db 721 | CTCTACTCGCGAGAGTTTCACCGAGTAGCCCTTGCCTGAGTACACAGTACAC 780 |
| Qy 781 | CTGGGGGACACTACGCTCGCGCAGCTATGGTGGGCTTCTGCTTCTGGCTCTCTCC 840 |
| Db 781 | CTGGGGGACACTACGCTCGCGCAGCTATGGTGGGCTTCTGCTTCTGGCTCTCTCC 840 |
| Qy 841 | AACGTGCTCTCTCAACGCGCGCCGCTCTACGAGGCTGCGCACTGCTACCAACCGGA 900 |
| Db 841 | AACGTGCTCTCTCAACGCGCGCCGCTCTACGAGGCTGCGCACTGCTACCAACCGGA 900 |
| Qy 901 | GCCTTGGCGCTCTTGGGGTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 960 |
| Db 901 | GCCTTGGCGCTCTTGGGGTCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 960 |

RESULT 2

| | |
|----------|--|
| ADD19228 | ADD19228 standard; cDNA; 1491 BP. |
| ID | ADD19228 standard; cDNA; 1491 BP. |
| XX | ADD19228; |
| AC | ADD19228; |
| DT | 15-JAN-2004 (first entry) |
| XX | Human cDNA from secreted protein gene 45. |
| DE | human secreted protein; cytostatic; antibacterial; virucide; |
| XX | neuroprotective; gynaecological; gastrointestinal; Gen: cardiant; |
| KW | cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen; |
| KW | respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic; |
| KW | nootropic; antiallergic; cancer; bacterial infection; viral infection; |
| KW | neural disorder; immune system disorder; blood disorder; |
| KW | muscular disorder; reproductive disorder; gastrointestinal disorder; |
| KW | pulmonary disorder; cardiovascular disorder; renal disorder; |
| KW | inflammatory disorder; proliferative disorder; human; ss; gene. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO2003052377-A2. |
| XX | |
| PD | 26-JUN-2003. |
| XX | |
| PF | 06-NOV-2002; 2002WO-US035606. |
| XX | |
| PR | 07-NOV-2001; 2001US-0331046P. |
| XX | |
| PA | (HUMA-) HUMAN GENOME SCI INC. |
| XX | |
| PI | Rosen CA, Ruben SM; |
| XX | |
| PI | WPI; 2003-533050/50. |
| XX | |
| DR | |

DR P-PSDB; ADD19303.
XX New isolated nucleic acids encoding signal transduction pathway component
PT polypeptides, useful for diagnosing, treating, and/or preventing
PT disorders, such as cancer, infections, cardiovascular and inflammatory
PT diseases.
XX Claim 1; SEQ ID NO 55; 554pp; English.
XX The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human secreted protein, representing one of 85 novel genes.
CC Also included are recombinant vectors, host cells (expressing the
CC protein), the secreted proteins (including their fragments, epitopes and
CC homologues), an isolated antibody that binds specifically to the protein,
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition (comprising determining the presence or absence of a mutation
CC in the nucleic acid and diagnosing a condition based on the presence or
CC absence of the mutation), diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or amount of expression of the protein in a biological sample
CC and diagnosing a condition based on the presence or amount of expression
CC of the protein), preventing, treating or ameliorating a medical condition
CC by administering the nucleic acid or protein to a mammalian subject,
CC identifying a binding partner to the protein, the gene corresponding to
CC the cDNA sequence, and identifying an activity in a biological assay
CC (comprising expressing the nucleic acid in a cell, isolating the
CC supernatant, detecting an activity in a biological assay and identifying
CC the protein in the supernatant having the activity). The nucleic acids
CC and proteins display the following activities: cytostatic, antibacterial,
CC virucide, neuroprotective, gynaecological, gastrointestinal, Gen.,
CC Cardiac, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-
CC Gen, Respiratory-Gen, Immunosuppressive, Cerebroprotective, Vastotropic,
CC Nootropic, Antiallergic. The methods and compositions of the present
CC invention are useful for diagnosing, treating, preventing and/or
CC prognosticating disorders related to the novel polypeptides, such as
CC cancer, bacterial or viral infections, and neural, immune system, blood,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, inflammatory or proliferative disorders (many examples of these
CC diseases and disorders are given in the specification). The present
CC sequence encodes a novel secreted protein of the invention.
XX
SQ Sequence 1491 BP; 293 A; 485 C; 410 G; 302 T; 0 U; 1 Other;
Query Match 99.7%; Score 1436; DB 9; Length 1491;
Best Local Similarity 99.9%; Pred. No. 1.7e-310;
Matches 1436; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
1 AAAGTAACGGCTACAGACAGTGGAGAAATAGTTGCTCGCCGGCTAGAAAACCTCTGTCG 60
44 AAAGTAACGGCTACAGACAGTGGAGAAATAGTTGCTCGCCGGCTAGAAAACCTCTGTCG 103
61 GTACCAACCCAGAGCGTTGAGAGAGAGCCACCTCCAGCTTCCTTAAACGAGAGGTGCA 120
104 GTACCAACCCAGAGCGTTGAGAGAGAGCCACCTCCAGCTTCCTTAAACGAGAGGTGCA 163
121 GGACTCAGACTTACAGACCCACTGGTCCAGCTTGTACGCAAGAGAGAGCCAGAGAC 180
164 GGACTCAGACTTACAGACCCACTGGTCCAGCTTGTACGCAAGAGAGAGCTCAAGGAC 223
181 GGCTCTCCCGGCTCCAGGAGCCAGCTTGTGCTTGGCTTGGCTTGGCTTGGCTTGGCT 240
224 GGCTCTCCCGGCTCCAGGAGCCAGCTTGTGCTTGGCTTGGCTTGGCTTGGCTTGGCT 283
241 ACTCGCCCGGCTGAGAGATGACCTTGGAGAGCGGTACTGCTTTTACCCAGGCC 300
284 ACTCGCCCGGCTGAGAGATGACCTTGGAGAGCGGTACTGCTTTTACCCAGGCC 343
301 CGGCATGCCAGCGCTTACAGCGTTCCACTGCTCATCTGTTAATCTAGTGTGTTTGGCTCTA 360
344 CGGCATGCCAGCGCTTACAGCGTTCCAGCGTTCCACTGCTCATCTGTTAATCTAGTGTGTTTGGCTCTA 403
361 GCAGCAGCTTCTGCTCATCTTGGCCGGGATCCGTTGGCCACTCGGCTGTGTTGGTGTG 420

DB 404 GCACCAAGCTTCTGCTCATCTTTCGCGGGGATCGTGGCCACTCGCGCTGTTGGTTG 463
QY 421 GTGAGAGTCTTCTCAGTCTGTTCATAGGCGCGAGAAATTTGGCTGTGCACTTCACTGCA 480
DB 464 GTGAGAGTCTTCTCAGTCTGTTCATAGGCGCGAGAAATTTGGCTGTGCACTTCACTGCA 523
QY 481 GAATGGTTCTGGGTACAGTGNACCAACACATCTTACAAAGCTTTCAGCGCAGCGCGC 540
DB 524 GAATGGTTCTGGGTACAGTGNACCAACACATCTTACAAAGCTTTCAGCGCAGCGCGC 583
QY 541 GTTACACCCGCTGTCTGCTGTGCTGTGCGGCTGTGAGGGCATTAATATATACATCAGAGG 600
DB 584 GTTACACCCGCTGTCTGCTGTGCTGTGCGGCTGTGAGGGCATTAATATATACATCAGAGG 643
QY 601 ACCCAGTGTGATCAGTGNACGAGACCATTTGACTTACAAAGCAGTGTTCACCTGCGCTG 660
DB 644 ACCCAGTGTGATCAGTGNACGAGACCATTTGACTTACAAAGCAGTGTTCACCTGCGCTG 703
QY 661 AAAGAGAAATACCGCGGAGTACGCGAAGCACTGTGAGAGAGGGCTGCGGACCCAGTGT 720
DB 704 AAAGAGAAATACCGCGGAGTACGCGAAGCACTGTGAGAGAGGGCTGCGGACCCAGTGT 763
QY 721 CTCTACCTGGGAGAGTTCACACCGAGTAGCCCTTGGCGCTGTACCCAGTGTACCCAGTGTAC 780
DB 764 CTCTACCTGGGAGAGTTCACACCGAGTAGCCCTTGGCGCTGTACCCAGTGTACCCAGTGTAC 823
QY 781 CTGCGCGGACACTACGCTCTGCGCACGCTATGGTGGGCTTCTGCTTCTTGCTCTCTCTCC 840
DB 824 CTGCGCGGACACTACGCTCTGCGCACGCTATGGTGGGCTTCTGCTTCTTGCTCTCTCTCC 883
QY 841 AACGTGCTCTTCCAGCGCGCGCGCTCTTACGGAGCCCTGGCAGCTGTGACACCGGA 900
DB 884 AACGTGCTCTTCCAGCGCGCGCGCTCTTACGGAGCCCTGGCAGCTGTGACACCGGA 943
QY 901 GCCTTCGCGCTCTTCCGCGCTCTTCCGCTTGGCTTCCATCTCTAGCGTGGCGCTCTGCGCG 960
DB 944 GCCTTCGCGCTCTTCCGCGCTCTTCCGCTTGGCTTCCATCTCTAGCGTGGCGCTCTGCGCG 1003
QY 961 CTCCGCTTACGCTCTCTGCGCTCTTCCGCTTGGCTTCCATCTCTAGCGTGGCGCTCTGCGCG 1020
DB 1004 CTCCGCTTACGCTCTCTGCGCTCTTCCGCTTGGCTTCCATCTCTAGCGTGGCGCTCTGCGCG 1063
QY 1021 GCAACCGCGCTCTCTGCTCTCTTCCGAGGGCGCTGGTGGTGTCTCCAGTATGTTGCG 1080
DB 1064 GCAACCGCGCTCTCTGCTCTCTTCCGAGGGCGCTGGTGGTGTCTCCAGTATGTTGCG 1123
QY 1081 CCCAGCGCTCTTCCGCTCTCTTCTGAGACCAAGCGCCAAAGGACTGACGACGAGAGAGG 1140
DB 1124 CCCAGCGCTCTTCCGCTCTCTTCTGAGACCAAGCGCCAAAGGACTGACGACGAGAGAGG 1183
QY 1141 GGCTCAGCTCTTATCTCTGCGGACCCACTGCAAGCAGCGCGCTCTTCCAGACTTAA 1200
DB 1184 GGCTCAGCTCTTATCTCTGCGGACCCACTGCAAGCAGCGCGCTCTTCCAGACTTAA 1243
QY 1201 TGTATCACCCTAACCTGTGAGGGGACCCCAATCTGGAATCTTTCGCGCTTGGGACAT 1260
DB 1244 TGTATCACCCTAACCTGTGAGGGGACCCCAATCTGGAATCTTTCGCGCTTGGGACAT 1303
QY 1261 CGCAGGCGGGAAGAGTGTGCGGACGCTTGGGCGCAGGAGCTCAGAGAGGGGACCTG 1320
DB 1304 CGCAGGCGGGAAGAGTGTGCGGACGCTTGGGCGCAGGAGCTCAGAGAGGGGACCTG 1363
QY 1321 AGCGCTGTGCGGAGGAGCTTGGGACATCCGAGCCACGAGGAAAGTCTCTGCGGCGGA 1380
DB 1364 AGCGCTGTGCGGAGGAGCTTGGGACATCCGAGGACACGAGGAAAGTCTCTGCGGCGGA 1423
QY 1381 TCTGTAATAAACCTTTTTTTTCTTTTCTTTTAAATAAAAAAAAAAAAAAAAAAAAAA 1438
DB 1424 TCTGTAATAAACCTTTTTTTTCTTTTCTTTTAAATAAAAAAAAAAAAAAAAAAAAAA 1481

RESULT 3
ABL90535

Db 1029 GCAACCGGCTCTGTGCTCTTCTCTCGAGGGGCGTGTGAGTCTCCAGTAATGTCG 1088
Qy 1081 CCAGGGCTCTTCCGACCTTCTTGGACCAAGGCCCAAGGACTGCGAGCCAGGAGAGG 1140
Db 1089 CCAGGGCTCTTCCGACCTTCTTGGACCAAGGCCCAAGGACTGCGAGCCAGGAGAGG 1148
Qy 1141 GGCTCACTCTTATCTCTGGGAGCCCACTGCAACAGCGCCCTCTCCAGACTTAAA 1200
Db 1149 GGCTCACTCTTATCTCTGGGAGCCCACTGCAACAGCGCCCTCTCCAGACTTAAA 1208
Qy 1201 TGTATCACCACTAACCTGTGAGGGGACCAATCTGACTCTCTCCCGCTTGGGACAT 1260
Db 1209 TGTATCACCACTAACCTGTGAGGGGACCAATCTGACTCTCTCCCGCTTGGGACAT 1268
Qy 1261 CGCAGCGCGGAGCAGTGGCCGCGAGCTGCGCCAGGAGAGCTCCAGAGGGGACATG 1320
Db 1269 CGCAGCGCGGAGCAGTGGCCGCGAGCTGCGCCAGGAGAGCTCCAGAGGGGACATG 1328
Qy 1321 AGCGCTGCTGGCGGAGGCTCGGACATCGCAGGACCAAGGAAAGTCTCTCTGGGCGA 1380
Db 1329 AGCGCTGCTGGCGGAGGCTCGGACATCGCAGGACCAAGGAAAGTCTCTCTGGGCGA 1388
Qy 1381 TCTGTAATAAATCTTTTCTTTTCTTTT 1413
Db 1389 TCTGTAATAAATCTTTTCTTTTCTTTT 1421

RESULT 5
AAC95564
ID AAC95564 standard; cDNA; 1460 BP.
AC AAC95564;
XX
DT 21-FEB-2001 (first entry)
DE Human secreted protein gene 44 SEQ ID NO:54.
XX
KW Human; secreted protein; cytostatic; immunosuppressive; nontropic;
KW neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
KW antiinflammatory; antitumor; anticonvulsant; antibacterial;
KW antifungal; antiparasitic; cardiant; cancer; immune disease; allergy;
KW cardiovascular disorder; wound healing; infection; neurological disease;
KW 85.
XX Homo sapiens.
XX WO20061596-A1.
PN 2000-611865/58.
XX 19-OCT-2000.
XX 06-APR-2000; 2000WO-US008983.
XX 09-APR-1999; 99US-0128703P.
PR 14-JAN-2000; 2000US-0176068P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
PI Ruben SM, Komatsoulis G;
XX WPI; 2000-611865/58.
DR P-PSDB; AAB52055.
XX
XX Fifty nucleic acid molecules encoding human secreted proteins, useful in
PT the prevention, treatment and diagnosis of cancer, immune disorders,
PT cardiovascular disorders and neurological diseases.
XX
XX Claim 1; Page 443; 505pp; English.
XX
XX Polynucleotide sequences AAC95521 - AAC95570 represent cDNA encoding 50
CC human secreted proteins AAB52012 - AAB52161. Sequences AAB52062 -
CC AAB52103 represent alternative polypeptides encoded by the genes, and

CC amino acid sequences with which they share homology. The genes and
CC proteins have activities dependent on the tissues and cells in which they
CC are expressed. Examples of their activities include cytostatic;
CC immunosuppressive; nontropic; neuroprotective; antiviral; antiallergic;
CC hepatotropic; antidiabetic; antiinflammatory; antitumor; anticonvulsant;
CC anticonvulsant; antibacterial; antifungal; antiparasitic; and cardiant.
CC The secreted proteins, polypeptides, antagonists and agonists may be
CC useful in treating, preventing and/or diagnosing diseases and disorders
CC such as cancer, particularly breast and ovarian cancer, and other cancers
CC of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract,
CC liver, lung, or urogenital. Immune disorders such as Addison's disease,
CC allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular diseases such as
CC myocardial ischaemias; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases such as viral,
CC bacterial, fungal and parasitic infections may also be treated using the
CC proteins and polynucleotides of the invention. Sequences AAC95512 -
CC AAC95520 and AAB52011 are used in the isolation and characterisation of
CC the proteins and polynucleotides of the invention
XX
SQ Sequence 1460 BP; 292 A; 472 C; 398 G; 298 T; 0 U; 0 Other;

Query Match 96.7%; Score 1392.8; DB 3; Length 1460;
Best Local Similarity 99.3%; Pred. No. 7.5e-301;
Matches 1430; Conservative 0; Mismatches 7; Indels 3; Gaps 3;
Qy 1 AAAGTAAAGGCTACAGACAGTGCAGAAATAGTTTCCTCGCGGCTAGAAAACTCTGCG 60
Db 22 AAAGTAAAGGCTACAGACAGTGCAGAAATAGTTTCCTCGCGGCTAGAAAACTCTGCG 81
Qy 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGTTCTTAAACGAGAGGTGCA 120
Db 82 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGTTCTTAAACGAGAGGTGCA 141
Qy 121 GGACTCAGACTTCACAGCCACTCGGTCCAGCCCTTACGAAAGAGAGCCCAAGGAC 180
Db 142 GGACTCAGACTTCACAGCCACTCGGTCCAGCCCTTACGAAAGAGAGCGTCAAGGAC 201
Qy 181 GGGCTCTCCGGTCCAGGAGCCCGACCTTGTGGTGGTGGCTGCGCGCTGCGTGCAGC 240
Db 202 GGGCTCTCCGGTCCAGGAGCCCGACCTTGTGGTGGTGGCTGCGCGCTGCGTGCAGC 261
Qy 241 ACTCGCGCGGTGCGAGCATGACCTGTGGAGCGGCTACTGCGCTTTTACCCCGAGCC 300
Db 262 ACTCGCGCGGTGCGAGCATGACCTGTGGAGCGGCTACTGCGCTTTTACCCCGAGCC 321
Qy 301 CGGCATGCGCAGGCTTCAGCGTTCCACTGCTCATGCTTATTTAGTGTGTTGCTCTA 360
Db 322 CGGCATGCGCAGGCTTCAGCGTTCCACTGCTCATGCTTATTTAGTGTGTTGCTCTA 381
Qy 361 CGAGCAAGCTTCTGCTCTCATCTGCGGGGATCGGTGGCCACTCGCGTGGTTTGGTGTG 420
Db 382 CGAGCAAGCTTCTGCTCTCATCTGCGGGGATCGGTGGCCACTCGCGTGGTTTGGTGTG 441
Qy 421 GTGAGAGTTCTTCTCAGTCTCTTCAATAGCGCGAGAAATGTGGGTGTGCACTTCACTGCA 480
Db 442 GTGAGAGTTCTTCTCAGTCTCTTCAATAGCGCGAGAAATGTGGGTGTGCACTTCACTGCA 501
Qy 481 GAATGGTGTGGGTACAGTGAACCAACATCTCAAGCCCTTCAGCGAGCGCGCG 540
Db 502 GAATGGTGTGGGTACAGTGAACCAACATCTCAAGCCCTTCAGCGAGCGCGCG 561
Qy 541 GTTACAGCCCGTGTGCTGCTCTGTTGGGCTCGAGGGCATTATATATACACTCAAGGG 600
Db 562 GTTACAGCCCGTGTGCTGCTGTTGGGCTCGAGGGCATTATATATACACTCAAGGG 621
Qy 601 ACCCGAGTGTGCTGAGTGAACAGAGCAATTTGACTATACAGAGAGCTTCACTTGGGCTGTG 660
Db 622 ACCCGAGTGTGCTGAGTGAACAGAGCAATTTGACTATACAGAGAGCTTCACTTGGGCTGTG 681
Qy 661 AAAGAGTAATACCGCGCGGAGTACCGAAGCAGCTGAGAGAGGGGCTCCCGAGCCAGTG 720

CC cancers, particularly ovarian or colon cancer, autoimmune diseases and
 CC related conditions
 XX
 SQ Sequence 498 BP; 97 A; 166 C; 137 G; 98 T; 0 U; 0 Other;
 Query Match 32.3%; Score 465.4; DB 3; Length 498;
 Best Local Similarity 99.4%; Pred. No. 4.4e-94;
 Matches 488; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
 QY 939 CTCTAGCGTCCGCTCTGCGCGCTCCGCTAGGCTCTCCGCGTCAACCACTCAGTAGC 997
 Db 1 CTCTAGCGTCCGCTCTGCGCGCTCCGCTAGGCTCTCCGCGTCAACCACTCAGTAGC 60
 QY 998 -CGCGCGCTCTGCGTCAAGCTGCGCAACCGGGGCTCTGCGCTCTCTCTCGAGGGGCGC 1056
 Db 61 AGCGCGCGCTCTGCGTCAAGCTGCGCAACCGGGGCTCTGCGCTCTCTCTCGAGGGGCGC 120
 QY 1057 GTGGTGAGTCTCCAGTATGTTGGCGCCAGCGCTCTTCGACACCTCTTCGACCAAGCGGC 1116
 Db 121 GTGGTGAGTCTCCAGTATGTTGGCGCCAGCGCTCTTCGACACCTCTTCGACCAAGCGGC 180
 QY 1117 AAGGACTGCGAGCAGAGAGAGGGGCTCACTCTTTATCTCTGGGACCCCACTGCACAAAG 1176
 Db 181 AAGGACTGCGAGCAGAGAGAGGGGCTCACTCTTTATCTCTGGGACCCCACTGCACAAAG 240
 QY 1177 CAGGCGGCTCTCCAGACTTAAATGTATCACCACTTAACCTGTGAGGGGACCCCAATCTG 1236
 Db 241 CAGGCGGCTCTCCAGACTTAAATGTATCACCACTTAACCTGTGAGGGGACCCCAATCTG 300
 QY 1237 GACTCTCTCCCGCTCTGGGACATCGAGCGGGGAGAGTGGCGGCGAGCGCTGGGCGC 1296
 Db 301 GACTCTCTCCCGCTCTGGGACATCGAGCGGGGAGAGTGGCGGCGAGCGCTGGGCGC 360
 QY 1297 AGGAGAGTCCAGGAGGACATGAGCGTGTGGCGGAGCGCTGGGCGAGCTGGGCGAGC 1356
 Db 361 AGGAGAGTCCAGGAGGACATGAGCGTGTGGCGGAGCGCTGGGCGAGCTGGGCGAGC 420
 QY 1357 ACCAGGGAAGTCTCTCTGGGCGATCTGTAAATAAACCTTTTCTTTTGTTTTAA 1416
 Db 421 ACCAGGGAAGTCTCTCTGGGCGATCTGTAAATAAACCTTTTCTTTTGTTTTAA 480
 QY 1417 AAAAAA 1427
 Db 481 AAAAAA 491
 RESULT 10
 ABZ68115
 ID ABZ68115 standard; DNA; 5033 BP.
 AC ABZ68115;
 DT 26-MAR-2003 (first entry)
 XX Human secreted protein encoding genomic DNA SEQ ID NO 1638.
 DE
 XX Human; secreted protein; nootropic; neuroprotective; cytostatic;
 KW virucide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
 KW vulvar; antibacterial; antiparkinsonian; antiskinning; antianaemic;
 KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
 KW antiinflammatory; antiallergic; antidiabetic; antidiabetic; anticonvulsant;
 KW antifungal; antiparasitic; cardiant; immune disorder; infection; vaccine;
 KW cardiovascular disorder; neurological disease; nephrotropic;
 KW gene therapy; gene; ds.
 XX Homo sapiens.
 OS
 XX WO20027186-A2.
 PN
 XX 03-OCT-2002.
 PD
 XX 26-MAR-2002; 2002WO-US009188.
 PF
 XX

PR 27-MAR-2001; 2001US-0278650P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Ruben SM;
 FI WPI; 2003-040583/03.
 DR
 XX New invention relates to novel human genes (ABZ68115-ABZ68209) and the
 CC encoded secreted proteins (ABP99470-ABP99872) useful for preventing,
 CC treating or ameliorating medical conditions e.g. by protein or gene
 CC therapy. The genes are isolated from a range of human tissues disclosed
 CC in the specification. The nucleic acids, proteins, antibodies and
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
 CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
 CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
 CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
 CC bacterial, fungal and parasitic infections
 XX
 SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;
 Query Match 28.1%; Score 405.4; DB 7; Length 5033;
 Best Local Similarity 98.6%; Pred. No. 2e-80;
 Matches 409; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 AAAGTAAACGGCTACAGACAGTGAAGATAGTTTCGTCGCGGCTAGAAAACTCTGTCG 60
 Db 13 AAAGTAAACGGCTACAGACAGTGAAGATAGTTTCGTCGCGGCTAGAAAACTCTGTCG 72
 QY 61 GTACCAACCCAGAGCGGTTGAGAGCAGCCACCTCCAGCTTCCTTAACGGAGAGGTGCA 120
 Db 73 GTACCAACCCAGAGCGGTTGAGAGCAGCCACCTCCAGCTTCCTTAACGGAGAGGTGCA 132
 QY 121 GGACTCAGACTTCACCAAGCCACTCGGTCCAGCGCTTGTACGCAAGAGAGCCCAAGGAC 180
 Db 133 GGACTCAGACTTCACCAAGCCACTCGGTCCAGCGCTTGTACGCAAGAGAGCCCAAGGAC 192
 QY 181 GCGCTCTCCCGCTCCAGGAGCCCGCCAGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCT 240
 Db 193 GCGCTCTCCCGCTCCAGGAGCCCGCCAGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCT 252
 QY 241 ACTGGCGCGGTTGAGAGTACCTCTGGAACCGGCTACTGCTTCCTTTTACCCAGGCC 300
 Db 253 ACTGGCGCGGTTGAGAGTACCTCTGGAACCGGCTACTGCTTCCTTTTACCCAGGCC 312
 QY 301 CGCATGCGCGGCTTCAGCGGTTCCACTGCTCATGCTTATCTAGTGTCTTTTGGCTCTA 360
 Db 313 CGCATGCGCGGCTTCAGCGGTTCCACTGCTCATGCTTATCTAGTGTCTTTTGGCTCTA 372
 QY 361 GCAGCAAGCTTCCTGCTCATCTTCCCGGGGATCGGTGCCCACTCGCGCTGCTTT 415
 Db 373 GCAGCAAGCTTCCTGCTCATCTTCCCGGGGATCGGTGCCCACTCGCGCTGCTTT 427
 RESULT 11
 ABZ74587
 ID ABZ74587 standard; DNA; 5033 BP.
 XX
 AC ABZ74587;

CC cardiovascular disorders (e.g. atherosclerosis and myocarditis);
CC infectious disease (e.g. HIV/AIDS); endocrine disorders (e.g. diabetes);
CC and gastrointestinal disorders (e.g. duodenal ulcers and
CC gastroenteritis). The present DNA sequence was used in the
CC exemplification of the invention.
XX
SQ Sequence 5033 BP; 1061 A; 1409 C; 1465 G; 1098 T; 0 U; 0 Other;

Query Match 28.1%; Score 405.4; DB 9; Length 5033;
Best Local Similarity 98.0%; Pred. No. 28-80;
Matches 409; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTGCGCGCTAGAAAACTCTGTGC 60
DB 13 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTGCGCGCTAGAAAACTCTGTGC 72
QY 61 GTACCAACCCCGAGAGCTTGAGAGCGCCACCTCCAGCTTCTTAACGGAGAGGTGCA 120
DB 73 GTACCAACCCCGAGAGCTTGAGAGCGCCACCTCCAGCTTCTTAACGGAGAGGTGCA 132
QY 121 GGAATCAGACTTCCAGAGCCCACTGGTCCAGGCTTGTACCAAGAGAGCCCAAGGAC 180
DB 133 GGAATCAGACTTCCAGAGCCCACTGGTCCAGGCTTGTACCAAGAGAGCCCAAGGAC 192
QY 181 GGGCTCTCCGGTCCAGAGCGCCAGCTTGTGCTGCTTGGCTCCGCTCGGTGCAGC 240
DB 193 GGGCTCTCCGGTCCAGAGCGCCAGCTTGTGCTGCTTGGCTCCGCTCGGTGCAGC 252
QY 241 ACTCGGCCGCGTGCAGATGACCTGTGGAACGGGTACTGCTTTTACCCCGAGGCC 300
DB 253 ACTCGGCCGCGTGCAGATGACCTGTGGAACGGGTACTGCTTTTACCCCGAGGCC 312
QY 301 CGGCATGCGCGAGGCTTCAGCGTTCACCTCATCTTATCTAGTGTGTTTGGCTCTA 360
DB 313 CGGCATGCGCGAGGCTTCAGCGTTCACCTCATCTTATCTAGTGTGTTTGGCTCTA 372
QY 361 GCAGCAGACTTCTGCTCATCTTCCGCGGATCCGTCGACCTCGCGCTGTTT 415
DB 373 GCAGCAGACTTCTGCTCATCTTCCGCGGATCCGTCGACCTCGCGCTGTTT 427

RESULT 13

ADD19187
ID ADD19187 standard; cDNA; 580 BP.

XX ADD19187;

AC 15-JAN-2004 (first entry)

DT Human cDNA from secreted protein gene 4.

XX human secreted protein; cytostatic; antibacterial; virucide;
XX neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW neutropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; human; ss; gene.

OS Homo sapiens.

XX W02003052377-A2.

PN 26-JUN-2003.

XX 06-NOV-2002; 2002WO-US035606.

XX 07-NOV-2001; 2001US-0331046P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX

PI

Rosen CA, Ruben SM;

XX WPI; 2003-533050/50.

DR P-FSDB; ADD19262.

XX

PT New isolated nucleic acids encoding signal transduction pathway component
PT polypeptides, useful for diagnosing, treating, and/or preventing
PT disorders, such as cancer, infections, cardiovascular and inflammatory
PT diseases.

XX Claim 1; SEQ ID NO 14; 554pp; English.

XX The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human secreted protein, representing one of 85 novel genes.
CC Also included are recombinant vectors, host cells (expressing the
CC protein), the secreted proteins (including their fragments, epitopes and
CC homologues), an isolated antibody that binds specifically to the protein,
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition (comprising determining the presence or absence of a mutation
CC in the nucleic acid and diagnosing a condition based on the presence or
CC absence of the mutation), diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or amount of expression of the protein in a biological sample
CC and diagnosing a condition based on the presence or amount of expression
CC of the protein), preventing, treating or ameliorating a medical condition
CC by administering the nucleic acid or protein to a mammalian subject, to
CC identifying a binding partner to the protein, the gene corresponding to
CC the cDNA sequence, and identifying an activity in a biological assay
CC (comprising expressing the nucleic acid in a cell, isolating the
CC supernatant, detecting an activity in a biological assay and identifying
CC the protein in the supernatant having the activity). The nucleic acids
CC and proteins display the following activities: Cytostatic, antibacterial,
CC Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen,
CC Cardiant, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-
CC Gen, Respiratory-Gen, Immunosuppressive, Cerebroprotective, Vasotropic,
CC Nootropic, Antiallergic. The methods and compositions of the present
CC invention are useful for diagnosing, treating, preventing and/or
CC prognosticating disorders related to the novel polypeptides, such as
CC cancer, bacterial or viral infections, and neural, immune system, blood,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, inflammatory or proliferative disorders (many examples of these
CC diseases and disorders are given in the specification). The present
CC sequence encodes a novel secreted protein of the invention.

XX Sequence 580 BP; 115 A; 198 C; 156 G; 111 T; 0 U; 0 Other;

Query Match 28.1%; Score 405; DB 9; Length 580;

Best Local Similarity 100.0%; Pred. No. 1.4e-80;

Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTGCGCGCTAGAAAACTCTGTGC 60

DB 13 AAAGTAACGGCTACAGACAGTGAAGAAATAGTTTGGCTGCGCGCTAGAAAACTCTGTGC 72

QY 61 GTACCAACCCCGAGAGCTTGAGAGCGCCACCTCCAGCTTCTTAACGGAGAGGTGCA 120

DB 73 GTACCAACCCCGAGAGCTTGAGAGCGCCACCTCCAGCTTCTTAACGGAGAGGTGCA 132

QY 121 GGAATCAGACTTCCAGAGCCCACTGGTCCAGGCTTGTACCAAGAGAGCCCAAGGAC 180

DB 133 GGAATCAGACTTCCAGAGCCCACTGGTCCAGGCTTGTACCAAGAGAGCCCAAGGAC 192

QY 181 GGGCTCTCCGGTCCAGAGCGCCAGCTTGTGCTGCTTGGCTCCGCTCGGTGCAGC 240

DB 193 GGGCTCTCCGGTCCAGAGCGCCAGCTTGTGCTGCTTGGCTCCGCTCGGTGCAGC 252

QY 241 ACTCGGCCGCGTGCAGATGACCTGTGGAACGGGTACTGCTTTTACCCCGAGGCC 300

DB 253 ACTCGGCCGCGTGCAGATGACCTGTGGAACGGGTACTGCTTTTACCCCGAGGCC 312

QY 301 CGGCATGCGCGAGGCTTCAGCGTTCACCTCATCTTATCTAGTGTGTTTGGCTCTA 360

DB 313 CGGCATGCGCGAGGCTTCAGCGTTCACCTCATCTTATCTAGTGTGTTTGGCTCTA 372

QY 361 GCAGCAAGCTTCTCTCTCATCTTGGCGGGATCCCTGGCCACTCG 405
 DB 373 GCAGCAAGCTTCTCTCTCATCTTGGCGGGATCCCTGGCCACTCG 417

RESULT 14
 AAS6563/c
 ID AAS6563 standard; cDNA; 406 BP.
 AC AAS6563;
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #2367.
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX Homo sapiens.
 XX WO200175067-A2.
 PN 11-OCT-2001.
 XX
 PD 30-MAR-2001; 2001WO-US008631.
 XX
 PF 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 DR P-PSDB; ABG02376.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensic, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 2367; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 406 BP; 83 A; 107 C; 138 G; 78 T; 0 U; 0 Other;
 Query Match 27.9%; Score 401.8; DB 5; Length 406;
 Best Local Similarity 99.5%; Pred. No. 6.4e-80;
 Matches 403; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 AAAGTAACGGCTACAGACAGTGAGAAATAGTTTCGTCGCGGCTAGAAAAACTCTGTGCG 60

Db 405 AAAGTAACGGCTACAGACAGTGAGAAATAGTTTCGTCGCGGCTAGAAAAACTCTGTGCG 346
 Qy 61 GTACCAACCCAGAGCGGTTGAGAGCAGCCACCTCCACGCTTCCTTAAACGAGAGGTGCA 120
 Db 345 GGACCAACCCAGAGCGGTTGAGAGCAGCCACCTCCACGCTTCCTTAAACGAGAGGTGCA 286
 Qy 121 GGACTCAGACTTTCACCAAGCCCACTCGCTCCAGCCTTGTATCGCAAGAGAGCGCAAGGAC 180
 Db 285 GGACTCAGACTTTCACCAAGCCCACTCGCTCCAGCCTTGTATCGCAAGAGAGCGTCAAGGAC 226
 Qy 181 GCGCTCTCCGCGTCCAGGAGCCCACTCGCTCCAGCCTTGTATCGCAAGAGAGCGTCAAGGAC 240
 Db 225 GCGCTCTCCGCGTCCAGGAGCCCACTCGCTCCAGCCTTGTATCGCTTCCGCTTCCGCTGAGC 166
 Qy 241 ACTCGGCGGCGTCCAGCATGACCCCTGTGAAACGCGCTACTGCGCTTTTATACCCCAAGCCC 300
 Db 165 ACTCGGCGGCGTCCAGCATGACCCCTGTGAAACGCGCTACTGCGCTTTTATACCCCAAGCCC 106
 Qy 301 CGGCATGCGGAGCGCTTCCAGGTTCCACTGCTCATGCTTATCTAGTGTGTTTGGCTCTA 360
 Db 105 CGGCATGCGGAGCGCTTCCAGGTTCCACTGCTCATGCTTATCTAGTGTGTTTGGCTCTA 46
 Qy 361 GCAGCAAGCTTCTCTCTCATCTTGGCGGGATCCCTGGCCACTCG 405
 Db 45 GCAGCAAGCTTCTCTCTCATCTTGGCGGGATCCCTGGCCACTCG 1

RESULT 15
 ADA53627/c
 ID ADA53627 standard; cDNA; 2684 BP.
 XX ADA53627;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human coding sequence, SEQ ID 1195.
 XX
 KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 XW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 XX inflammatory disease; osteoporosis; neurological disease; gene; ss.
 OS Homo sapiens.
 OS
 PN EP1293569-A2.
 XX
 PD 19-MAR-2003.
 XX
 PF 21-MAR-2002; 2002EP-00006586.
 XX
 PR 14-SEP-2001; 2001JP-00328381.
 XX
 PR 24-JAN-2002; 2002US-0350435P.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;
 XX
 DR WPI; 2003-395539/38.
 DR P-PSDB; ADA55266.
 XX
 XX New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.
 XX
 PS Claim 1; SEQ ID NO 1195; 205pp; English.
 XX
 CC The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of
 CC diseases caused by abnormalities of the proteins, e.g. cancer,

CC inflammatory diseases, osteoporosis or neurological disease.

XX Sequence 2684 BP; 599 A; 785 C; 684 G; 616 T; 0 U; 0 Other;

Query Match 27.5%; Score 396; DB 7; Length 2684;
Best Local Similarity 97.6%; Pred.No. 2.1e-78;
Matches 402; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

| | | | | | | | | | |
|----|------|-----------|----------|----------|----------|------------|----------|--------|------|
| QY | 1007 | TCTGGGTCA | CGCTGGCA | ACCGGCTC | CTGCTCTT | CTCTCGAGGG | GGCCGTTG | AGTTC | 1066 |
| Db | 2363 | TTTCGATC | CCCCAC | CGCCAC | AGGGTCTG | CTCTCTTC | CTCGAGGG | CCGTTG | 2304 |
| QY | 1067 | TCCAGTAT | TTTCGG | CCAGCG | CTCTTC | CGACCC | TTCTG | ACCAAG | 1126 |
| Db | 2303 | TCCAGTAT | TTTCGG | CCAGCG | CTCTTC | CGACCC | TTCTG | ACCAAG | 2244 |
| QY | 1127 | GCCAGAG | AGAGGG | GCTCAC | CTTATC | CTCGGG | ACCCAC | TGCAAG | 1186 |
| Db | 2243 | GCCAGAG | AGAGGG | GCTCAC | CTTATC | CTCGGG | ACCCAC | TGCAAG | 2184 |
| QY | 1187 | TCCAGAC | TAAATG | TATAC | CACTAA | CTGTAG | GGGAC | CCCAAT | 1246 |
| Db | 2183 | TCCAGAC | TAAATG | TATAC | CACTAA | CTGTAG | GGGAC | CCCAAT | 2124 |
| QY | 1247 | CGCCCTT | GGACAT | CGAG | CGGG | AGCAG | TGCG | CCGAG | 1306 |
| Db | 2123 | CGCCCTT | GGACAT | CGAG | CGGG | AGCAG | TGCG | CCGAG | 2064 |
| QY | 1307 | CAGGAAG | GGCCTG | AGCG | CTGT | GGCG | CGAG | GCTCG | 1366 |
| Db | 2063 | CAGGAAG | GGCCTG | AGCG | CTGT | GGCG | CGAG | GCTCG | 2004 |
| QY | 1367 | GTCTCTT | GGGCGA | CTGT | TAAT | TAAC | CTTTT | CTTTT | 1418 |
| Db | 2003 | GTCTCTT | GGGCGA | CTGT | TAAT | TAAC | CTTTT | CTTTT | 1952 |

Search completed: June 3, 2004, 10:20:49
Job time : 646 secs

Best Local Similarity 1.9%; Pred. No. 5.5e-05;
Matches 7; Conservative 221; Mismatches 131; Indels 0; Gaps 0;
QY 753 CCCTTCGGCGCTTACACAGTACACCTGGCGGACACTAGCTCGCCCAAGCTATG 812
Db 1084 YY 1143
QY 813 GGTGGGCTTCTGCTTCTGCTCTCCCAAGTCTCTCCACGCGCGCGCTCTA 872
Db 1144 YY 1203
QY 873 CGAGGCTGGCACTGTACACCGAGCTTCGCGCTTCGGGGTCTTCGCTTGGC 932
Db 1204 YY 1263
QY 933 CTCATCTCTAGCTGCTGCGCTTCGCGCTTCGCTAGGCTCTCCGCGCTACCACTA 992
Db 1264 YY 1323
QY 993 GTAGGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1052
Db 1324 YY 1383
QY 1053 GCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1111
Db 1384 YY 1442

RESULT 2
US-09-252-991A-1304/c
; Sequence 1304, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1304
; LENGTH: 1098
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1304

Query Match 3.5%; Score 50.2; DB 4; Length 1098;
Best Local Similarity 50.6%; Pred. No. 0.011;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 830 GGCTCTCTCCAACTGCTGCTCTCCACGCGCGCGCTCTACGAGGCGCTGCACTGC 889
Db 602 GACTCTCTCAATGCTGCTGCGCGGCTGCGGAGCTGCTGCGGCGCGCTGATGCTGC 543
QY 890 TGACCAACGAGCTTCGCGCTCTTCGCGGCTCTTCGCTTCGCTTCATCTAGCGTGC 949
Db 542 TGGCGCGCGCTGCGGCTGCGGCTGCGGAGCTGCTGCGCTGCGCTGCGGCGCG 483
QY 950 CGCTCTGCGCGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1009
Db 482 TGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
QY 1010 GGTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
Db 422 TGGCGCTGCGCGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 164

RESULT 3
US-09-252-991A-1255

; Sequence 1255, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1255
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1255

Query Match 3.5%; Score 50.2; DB 4; Length 1257;
Best Local Similarity 50.6%; Pred. No. 0.011;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 830 GGCTCTCTCCAACTGCTGCTCTCCACGCGCGCGCTCTACGAGGCGCTGCACTGC 889
Db 333 GACTCTCTCAATGCTGCTGCGCGGCTGCGGAGCTGCTGCGGCGCGCTGATGCTGC 392
QY 890 TGACCAACGAGCTTCGCGCTCTTCGCGGCTCTTCGCTTCGCTTCATCTAGCGTGC 949
Db 393 TGGCGCGCGCTGCGGCTGCGGCTGCGGAGCTGCTGCGCTGCGCTGCGGCGCG 452
QY 950 CGCTCTGCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCTTCGCGCT 1009
Db 453 TGCTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512
QY 1010 GGTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1068
Db 513 TGGCGCTGCGCGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCT 571

RESULT 4
US-09-252-991A-1163
; Sequence 1163, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1163
; LENGTH: 1746
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1163

Query Match 3.5%; Score 50.2; DB 4; Length 1746;
Best Local Similarity 50.6%; Pred. No. 0.013;
Matches 121; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
QY 830 GGCTCTCTCCAACTGCTGCTCTCCACGCGCGCGCTCTACGAGGCGCTGCACTGC 889
Db 1283 GACTCTCTCAATGCTGCTGCGCGGCTGCGGAGCTGCTGCGGCGCGCTGATGCTGC 1342
QY 890 TGACCAACGAGCTTCGCGCTCTTCGCGGCTCTTCGCTTCGCTTCATCTAGCGTGC 949

Db 1343 TGGGGCGGGCTGGGGCGAAGCGGGGAGCCCTCGCCCTCGCGCTGCACACCGGGCGG 1402
Qy 950 CGCTCTCGCGCTCGCGCTAGGCTCTCTCGCGCTCACCCTCAGTACGGCGCGCTTCT 1009
Db 1403 TGGTGGCGGGCTGTTGGCGAAGCCCTCGAGAACACCCCGGCGCAGCGCGAGGCT 1462
Qy 1010 GGGTCAGCTGGCAACCGGCTCTCTCTCTCTCGAGGGGCGGTGAGTCTC 1068
Db 1463 TGGCCTTGGCGGCGAGCGGCTCGCGGCTATGGCCTATGGCCTCGCCGGGGGTC 1521

RESULT 5
US-09-252-991A-1350/c
; Sequence 1350, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1350
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1350

Query Match 3.3%; Score 48.2; DB 4; Length 396;
Best Local Similarity 52.2%; Pred. No. 0.024;
Matches 107; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
Qy 830 GGCTCTCTCCACGCTGCTCTCCACCGCGCGCGCTCTACGAGCGCTGGCACTGC 889
Db 217 GACTCTCTCAAGCTGCTCGCGCGCGCTCGCGAGCTGTCTGGCGCGCGCTATGGTGC 158
Qy 890 TGACCAACCGAGCTTTCGCGCTCTTCGCGGCTCTCGCTTGGCTTGCCTTCTCTAGCGTGC 949
Db 157 TGGCGCGCGGGCTCGGCGCGCGAGCCCTCGCCCTGGCGCTGCACACCGCGCGG 98
Qy 950 CGCTCTCGCGCTCGCGCTAGGCTCTCTCGCGCTCAGCACTCAGTACGGCGCGCTTCT 1009
Db 97 TGGTGGCGGGCTGTTGGCGAAGCCCTCGAGAACACCCCGGCGCAGCGCGGCGCT 38
Qy 1010 GGGTCAGCTGGCAACCGGCGCTCT 1034
Db 37 TGGCGCTGGCGCGGAGCGGCT 13

RESULT 6
US-09-614-912-35
; Sequence 35, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Guo-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Caimi, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/170,906
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 60/172,959
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/172,946
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 35
; LENGTH: 1436
; TYPE: DNA
; ORGANISM: Zea mays
US-09-614-912-35
Query Match 3.3%; Score 48; DB 4; Length 1436;
Best Local Similarity 46.8%; Pred. No. 0.039;
Matches 187; Conservative 0; Mismatches 210; Indels 3; Gaps 1;
Qy 541 GTTACAGCCCGTCTCGTCTGTCTGTGGGCTTGAGGGGCAATTAATTACACTCAGGG 600
Db 43 GCTTCT 102
Qy 601 ACCCCAGTGCATCAGCTGAGCAAGAGACCATTCATCAACAGCAGTTCACCTGGCGCTCTG 660
Db 103 TTGACGCTCCCGCTTGAGCGGCTCTCTCGAATCCCGCAGCATTCATCTGGCGGG 162
Qy 661 AAAGAGAAATTAACCCCGGAGTACGCGAAGCGCACTGAGAGAGGGGCTGCGGACCCAGTG 720
Db 163 GACGAGAGCCCAACCCCGGAGCTCCCGCGAGAGCTGGCGCTGCGCTCATCGACTCTCTCC 222
Qy 721 CTCTACTCGCGGAGAGTTCACACCG---AGTAGCCCTTGGCGGCTGTACACCACTAGTAC 777
Db 223 GGGAGACCCCGCGAGTGTTCGCGAGGTCTCGCGCGCTGCGACTGACGCGCTTCTTC 282
Qy 778 CACTCTCGCGGACACTACGCTCTGCGGCGCTATGGTGGCGCTTCTGCTTCTGCTCTCTC 837
Db 283 CAGTGTGTGGGCAACCGCATCGACGCGCGGCTGACGCGGAGGCGCCACCGCTGATGGAC 342
Qy 838 TCCAAAGTGTCTCTCCACGCGCGCGCTCTACGAGGGCTTGCACCTGCTGACCACTC 897
Db 343 GCCTTCTTACGCTGCGCTCTCCCGGCAAGCAGCGCGCGCGCGCGCGCGCGCGGACAGC 402
Qy 898 GGAGCTTTCGCGCTCTTCGCGGCTTTCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTC 937
Db 403 TGGCGCTTACCGCAGCAGCTTACGCGGCGGCTTTCGCGCTTTCGCGCTTTCGCGCTTTC 442

RESULT 7
US-09-252-991A-11029/c
; Sequence 11029, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11029
; LENGTH: 420
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-11029

| Query Match | 3.38; | Score 47.6; | DB 4; | Length 420; |
|-----------------------|-----------------|--|-----------|-------------|
| Best Local Similarity | 48.2%; | Pred. No. 0.033; | | |
| Matches 134; | Conservative 0; | Mismatches 144; | Indels 0; | Gaps 0; |
| Qy | 785 | CGGACACTACGCTTCGGCCACGCTATGGTGGGTTCTCTCTGTGCTCTCTCTCAACG | 844 | |
| Db | 382 | CGGTGGAAAACGCTTCTCTCATGCCCTGGCTGGTCGGCACTGCACCTGATCCACTGCGTGG | 323 | |
| Qy | 845 | TGCTGCTCTCCACGCGCGCCCGCTCTACGGAGGCGCTGGCACTCTGACCAACGGAGCGCT | 904 | |
| Db | 322 | CGGTGACCGAGAACGGCGCGGTGTTCAAGACTGGACGGTGTCTTGGCGATCGCCGCTT | 263 | |
| Qy | 905 | TCGCGCTCTTCGGGGTCTTCGCGTTGAGCTTCATCTCTTAGCGTGCGCGCTCTGCGCGCTCC | 964 | |
| Db | 262 | TCTCCCTGAGCCTGCTGGGCACTTCTCTGTGTCGCTCCGGGTCGCTCACCTCGGTGCAAG | 203 | |
| Qy | 965 | GCCTAGACTCTTCGCGCTCACCACTCAGTACGGCGCGCTTCTTGCGTCACTGCTGGCAA | 1024 | |
| Db | 202 | CTTTCGGCTTCGACCGGAGCGCGGGGTGTTTCATCTCGCCCTTCTGCTGCTGGTGGTCG | 143 | |
| Qy | 1025 | CCGCGCTCTGTGCTCTTCTTCGAGAGGGGCGGTGGT | 1062 | |
| Db | 142 | CGCGCTCGCTGACCTGTTTCGCGCTTCGCGCTTCGCGCGCGCGGTG | 105 | |

RESULT 8

```

US-09-252-991A-11226
; Sequence 11226, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11226
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-11226

```

Query Match 3.3%: Score 47.6; DB 4; Length 1683;

| | Query Match | Best Local Similarity | 48.2%; | Pred. No. 0.051; | Mismatches 144; | Indels 0; | Gaps 0; |
|----|--------------|--|--------|------------------|-----------------|-----------|---------|
| | Matches 134; | Conservative | | | | | |
| QY | 785 | CGGACACTACGCGCTCGGCCACGGCTATGGGCGGGTTCTGCTTCTGGCTCTCTTCCAACG | 844 | | | | |
| DB | 864 | CGGTGGAAAAGCTCTTCATGCCCTGGCTGGTGGCACTGCATCATCCACTCGCTGG | 923 | | | | |
| QY | 845 | TCTGCTCTCCACGCCGCCCGCTCTACGGAGGCTGGCACTGCTGACACACGGAGCCT | 904 | | | | |
| DB | 924 | CGGTACCCAGAAAGCGCGCGCTGTTCAGAGCTGACCGTGTCTTCTGGCGATGCCCGCT | 983 | | | | |
| QY | 905 | TCCGCTCTTGGGGTCTTCGCTTGGCTCCATCTCTAGCGTGCCTGCTGCGCGCTCC | 964 | | | | |
| DB | 984 | TCTCCCTGAGCTGCTGGCACCTTCCTGCTCGCTCCGGGCTGCTCACTCGTGCACG | 1043 | | | | |
| QY | 965 | GCTAGGCTCTCCCGCTCTACCACCTCACTAGTAGCGCGCGCTTCTGGGTCACTGGCAA | 1024 | | | | |
| DB | 1044 | CCTTCGCTCCGACCGCGAGCGGGGTTCATCTCGCCTTCTGCTGCTGGTGTCTG | 1103 | | | | |
| QY | 1025 | CGGGCTCTGTGCTCTTCTCTGGAGGGCGGGTGGTG | 1062 | | | | |
| DB | 1104 | GGGGTCGTGATACCTGTTTGGCCCTTCGCGCGCGCGGTG | 1141 | | | | |

RESULT 9

```

US-09-252-991A-11111
; Sequence 11111, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 11111
; LENGTH: 1974
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-11111

```

Query Match 3.3%: Score 47.6; DB 4; Length 1974;

| | Query: NC_010422.1; 100.00%; 1000bp | Best Local Similarity 48.2%; Pred No. 0.054; | Mismatches 144; | Indels 0; | Gaps 0; |
|----|-------------------------------------|--|-----------------|-----------|---------|
| | Matches 134; | Conservative 0; | | | |
| Qy | 785 | CGGACACTACGGCTCGGGCACACCTATAGGTGGCGTTCTGCTTCTGGCTCCTCTCCACG | 844 | | |
| Db | 731 | CGGTGGAAAACGTTCTCTTCAATCCCTGGCTGTGTGGCACTGCACCTATTCACCTGCTGG | 790 | | |
| Qy | 845 | TGCTGCTCTCCACGGCGGGCCCGCTCTACGGAGGCCCTGGCACTGCTGTACACACGGAGCCT | 904 | | |
| Db | 791 | CGGTGACCGAAGAGCGCGCGTGTTCAGAGCTGGACCGTGCTTCTGGCGATCGCGCTT | 850 | | |
| Qy | 905 | TGGGCTCTTCGGGGTCTTCGGCTGGCGCTCCATCTCTAGCGTGGCGCTCTGCGCGTCC | 964 | | |
| Db | 851 | TCCTCCCTGAGCTGCTGGGCACCTTCCTGCTCGCTCCGGGGTCTCACCTCGGTGCACG | 910 | | |
| Qy | 965 | GCTTAGGCTCCTTCGGGCTCACCACTCAGTAGCGGCGCGGCTTCTGGTCAAGCTGGCAA | 1024 | | |
| Db | 911 | CCTTCGCTCCGACCCGAGCGGGGTGTTCATCTCGGCTCTCTCTGCTGCTGTGTCTG | 970 | | |
| Qy | 1025 | CCGGCGTCTGTGCTCTTCTCTCGAGGGGCCGTGGTG | 1062 | | |
| Db | 971 | GGGCTCGCTGACCTCTTCGCTGGCGCGCGGTG | 1008 | | |

RESULT 30

```

US-09-252-991A-10956/c
; Sequence 10956, Application US/09252991A
; Patent No. 6531795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10956
; LENGTH: 2283
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-10956

```

Query Match 3.3%; Score 47.6; DB 4; Length 2283;

| | | | |
|---------------------------|-------|--------------------|-----------|
| Query Match | 2.8% | | |
| Best Local Similarity | 48.2% | Pred. No. 0.056; | |
| Matches 134; Conservative | | 0; Mismatches 144; | Indels 0; |
| Matches 134; Conservative | | 0; Mismatches 144; | Indels 0; |
| Matches 134; Conservative | | 0; Mismatches 144; | Indels 0; |

785 CGGACACTACGCTCGCCACGCTATGATGGTGGGCTTCTGCTTCTGGCTCTCTCCACAG 844
2081 CGGTGGAAACGCTTCTTATGCTGCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTG 2022
845 TGCTGCTCTCACCGCCGCGCTCTACCGAGGCTTGGCTGCTGCTGCTGCTGCTGCTG 904
2021 CGGTGACCGAAGCGCGGCTGCTCAAGAGCTGGACCGGCTTCTGCTGCTGCTGCTG 1962
905 TCGGCTCTTTCGGGCTCTTTCGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 964
1961 TCTCCCTGAGCTGCTGGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1902
965 GCCTAGGCTCTTTCGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1024
1901 CCTTGGCTTTCGGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1842
1025 CCGGCTGCTGCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
1841 GCGGCTGCTGCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1804

RESULT 11

US-09-614-912-49
; Sequence 49, Application US/09614912
; Patent No. 667502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Gou-Hau
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Cai, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BE1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/143,412
; PRIOR FILING DATE: 1999-07-12
; PRIOR APPLICATION NUMBER: 60/146,650
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 60/170,906
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: 60/172,959
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 60/172,946
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 49
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1309)
; NAME/KEY: unsure
; LOCATION: (1339)
; NAME/KEY: unsure
; LOCATION: (1351)
; NAME/KEY: unsure
; LOCATION: (1402)
; NAME/KEY: unsure
; LOCATION: (1429)
US-09-614-912-49

Query Match

3.2%; Score 46.4; DB 4; Length 1431;

Best Local Similarity 46.5%; Pred. No. 0.094;
Matches 186; Conservative 0; Mismatches 211; Indels 3; Gaps 1;
QY 541 GTTACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
DB 33 GCTTCT 92
QY 601 ACCCCAGTGCATCAGCTGAAACGAGACCATTTGACTATACAGAGAGTTCACCTG 660
DB 93 TTCAGCGCTGCGCGCTGAGCGGCTCTCCGACATCCCGCAGCAGTTCATCTGCGCG 152
QY 661 AAGAGAGTAAGCCGCGGAGTAGCGGACGAGCAGTGGAGAGGGGCTGCGGACCCAG 720
DB 153 GACGAGAGCCCGGAGTCTCGCGGAGAGTCTGCGGAGAGTCTGCGGAGTCTGCGG 212
QY 721 CTCTACCTGCGGAGAGTTCACACCG---AGTAGCCCTTTCGGGCTGTACACCACT 777
DB 213 GGGGAAGCGCGAGGTGCTCGGAGTCTCGGAGTCTCGGAGTCTCGGAGTCTCGG 272
QY 778 CACCTGGGGGACACTAGCGCTCGGCGACGCTATGATGGTGGGCTTCTGCTTCTGCT 837
DB 273 CAGGTGGTGGGACGCGCATGACCGCGGCTGACGCGGAGGCGCCACCGCTGATGG 332
QY 838 TCCAAAGTGTCTCTCCACCGCGGCGGCTCTTACGAGGCTTGGGCTGTGACCAACC 897
DB 333 GCCTTCTTACGCTGCGCTCCCGGACAGCAGCGCGCGCGGCGGCGGAGGAGCAGC 392
QY 898 GGAGCGCTGCGGCTTTCGGGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCG 937
DB 393 TCGGCTACGCGCAGCTTTCACGCGGCGGCTTTCGCGCTTTCGCGCTTTCGCG 432

RESULT 12

US-09-249-585A-4
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0D905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match

Best Local Similarity 44.1%; Pred. No. 0.22;
Matches 189; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 825 CTCTGGCTCTCTCCAAAGTGTGCTCTCCACCGCGGCGGCTCTAGCGAGGCTTGGC 884
DB 706 CTTGCT 765
QY 885 ACTGCTGACACCGAGCTTTCGGCTTCTTTCGGGCTTTCGGCTTTCGGCTTTCGG 944
DB 766 CTTCT 825
QY 945 CGTGGCGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGGCTTTCGG 1004
DB 826 CTTCT 885
QY 1005 CTTCTGGGTACGCTGGGAAACCGGCTTCTTTCGGCTTCTTTCGGAGGGGCTTGG 1064
DB 886 CTTCT 945

QY 1065 TCTCCAGTATGTTGGCCGAGCGCTCTTCGACCCCTTCTTGAGACCAAGCGCCAAAGACTG 1124
Db 946 CCGCGTCTCTGCTCCCGCTCTCTGCTCTCCACCTCCGCGCCCGACCTCTCCGTCACCT 1005
QY 1125 CAGCCAGGAGAGGGGGCTCACTCTTATCTTCGCGGACCCACTGCACAAAGCGCGCG 1184
Db 1006 CCGGCCCGAGCTCTCATACCTCCGCGCCGAGCTCTCCATCCTCCGCGGCCCA 1065
QY 1185 TCTCCAGACTTAAATATGATCACTAACTGAGGGGAGCCAAATCTGAGCTCTT 1244
Db 1066 TCTCTGCACTTCTCGGTCCCGCCCTTCAGCACTTCTCGGTCCCGCTCTCCAGCACT 1125
QY 1245 CCGCGCTT 1253
Db 1126 GCACCTCTT 1134

RESULT 13
US-09-130-114-2
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

Query Match 3.1%; Score 45; DB 2; Length 1931;
Best Local Similarity 44.1%; Pred. No. 0.22;
Matches 183; Conservative 0; Mismatches 240; Indels 0; Gaps 0;

QY 825 CTTCTGGCTCTCTCCAGCGCTCTGCTCTTCGGGGCTTCGCGCTTGGCGCTCCATCTCTAG 884
Db 706 CTTCTGGCTCTCTCCAGCGCTCTGCTCTTCGGGGCTTCGCGCTTGGCGCTCCATCTCTAG 884
QY 885 ACTGCTGACACCGAGCGCTTCTGCTCTTCGGGGCTTCGCGCTTGGCGCTCCATCTCTAG 944
Db 766 CTTCTGGCTCTCTCCAGCGCTCTGCTCTTCGGGGCTTCGCGCTTGGCGCTCCATCTCTAG 825
QY 945 CTTCTGGCTCTCTCCAGCGCTTCTGCTCTTCGGGGCTTCGCGCTTGGCGCTCCATCTCTAG 1004
Db 826 CTTCTGGCTCTCTCCAGCGCTTCTGCTCTTCGGGGCTTCGCGCTTGGCGCTCCATCTCTAG 885
QY 1005 CTTCTGGCTCTCTCCAGCGCTTCTGCTCTTCGGGGCTTCGCGCTTGGCGCTCCATCTCTAG 1064
Db 886 CTTCTGGCTCTCTCCAGCGCTTCTGCTCTTCGGGGCTTCGCGCTTGGCGCTCCATCTCTAG 945
QY 1065 TCTCCAGTATGTTGGCCGAGCGCTCTTCGACCCCTTCTGAGCAAGCGCCAAAGACTG 1124
Db 946 CCGCGTCTCTGCTCCCGCTCTCTGCTCTCCAGCTCCGCGCCCGACCTCTCCGTCACCT 1005
QY 1125 CAGCCAGGAGAGGGGGCTCACTCTTATCTTCGCGGACCCACTGCACAAAGCGCGCG 1184
Db 1006 CCGGCCCGAGCTCTCATACCTCCGCGCCGAGCTCTCCATCCTCCGCGGCCCA 1065
QY 1185 TCTCCAGACTTAAATATGATCACTAACTGAGGGGAGCCAAATCTGAGCTCTT 1244
Db 1066 TCTCTGCACTTCTCGGTCCCGCCCTTCAGCACTTCTCGGTCCCGCTCTCCAGCACT 1125
QY 1245 CCGCGCTT 1253

Db 1126 GCACCTCTT 1134

RESULT 14
US-09-485-549-1
; Sequence 1, Application US/09485549
; Patent No. 6361948
; GENERAL INFORMATION:
; APPLICANT: James Tricoli
; APPLICANT: Rachel Rhondinelli
; APPLICANT: Fox Chase Cancer Center
; TITLE OF INVENTION: Prognostic Compositions for Prostate Cancer and Methods of Use
; TITLE OF INVENTION: thereof
; FILE REFERENCE: PCCC 96-13
; CURRENT APPLICATION NUMBER: US/09/485,549
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: PCT/US98/16768
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/055,285
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1835
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-485-549-1

Query Match 3.1%; Score 44.4; DB 4; Length 1835;
Best Local Similarity 59.5%; Pred. No. 0.3;
Matches 75; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 1316 CACTGAGCGCTCTGCGCGGAGCGCTCGACATCGCGAGCACCAGGAAAGTCTCCTGG 1375
Db 1658 CACTTTCGCGCGCTCTGCGCGGAGCGCTCGACATCGCGAGGATCGGCGATCGGAGG 1717
QY 1376 GCGGATCTGTAATAAACCTTTTCTTTTAAATAAAAAAAAAAAAAAAAAAAAA 1435
Db 1718 GCGGCTCTGAGATTAAAGATTTTACCTTTGGGTAAATAAAAAAAAAAAAAAAAAAAAA 1777
QY 1436 AAAAAA 1441
Db 1778 AAAAAA 1783

RESULT 15
US-09-128-155-16
; Sequence 16, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(152331)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

Query Match 3.1%; Score 44.4; DB 3; Length 152331;

| Best local Similarity 51.5%; Pred. No. 1.2; | |
|--|--|
| Matches 102; Conservative 0; Mismatches 96; Indels 0; Gaps 0; | |
| Qy 832 CTCTCTTCCACGTCCTCTTTCACGCGCGCGCGCTCTACGAGGCTTGGCTCTCTG 891 | |
| Db 22095 CCG 22154 | |
| Qy 892 ACCACCGGAGCTTCGCGCTCTTCGGGGTCTTCGGCTTGGCTTCATCTTAGCGTGCCG 951 | |
| Db 22155 ACCCG 22214 | |
| Qy 952 CTCTGCGCGCTTCGCGCTAGGCTCTTCGCGCTTCACACTCAGTACGGCGCGCTTCTGG 1011 | |
| Db 22215 CCGCACCG 22274 | |
| Qy 1012 GTCACGCTGGCAACCGGC 1029 | |
| Db 22275 CCG 22292 | |

Search completed: June 3, 2004, 13:07:59
Job time : 127 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 10:08:49 ; Search time 653 seconds
(without alignments)
10087.131 Million cell updates/sec

Title: US-09-936-456-1

Perfect score: 1441

Sequence: 1 aaagtaacgctacagacag.....aaaaaaaaaaaaaaaaaaaa 1441

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2995936 seqs, 228098010 residues

Total number of hits satisfying chosen parameters: 5991872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 1435.6 | 99.6 | 1474 | 16 | US-10-264-237-1097 |
| 2 | 1413 | 98.1 | 1421 | 15 | US-10-187-657-2 |
| 3 | 1407.8 | 97.7 | 1420 | 15 | US-10-187-657-4 |
| 4 | 694.6 | 48.2 | 1594 | 15 | US-10-187-657-9 |
| 5 | 451.6 | 31.3 | 522 | 15 | US-10-187-657-5 |
| 6 | 396 | 27.5 | 2684 | 16 | US-10-094-749-1195 |
| 7 | 385.8 | 26.8 | 450 | 15 | US-10-187-657-7 |
| 8 | 354.8 | 24.6 | 1029 | 10 | US-09-759-1308-424 |
| 9 | 354.8 | 24.6 | 1029 | 14 | US-10-042-431-54 |
| 10 | 354.8 | 24.6 | 2133 | 10 | US-09-759-1308-423 |
| 11 | 354.8 | 24.6 | 2133 | 14 | US-10-042-431-53 |
| 12 | 322 | 22.3 | 506 | 15 | US-10-187-657-6 |
| 13 | 300 | 20.8 | 346 | 15 | US-10-187-657-3 |
| 14 | 245.8 | 17.1 | 2684 | 16 | US-10-094-749-1195 |

| | | | | | | |
|----|-------|------|---------|----|---------------------|-------------------|
| 15 | 200.6 | 13.9 | 439 | 13 | US-10-027-632-91970 | Sequence 91970, A |
| 16 | 200.6 | 13.9 | 439 | 16 | US-10-027-632-91970 | Sequence 91970, A |
| 17 | 136.2 | 13.6 | 232 | 9 | US-09-783-590-9492 | Sequence 9492, Ap |
| 18 | 170 | 11.8 | 250 | 15 | US-10-187-657-8 | Sequence 8, Appl |
| 19 | 64 | 4.4 | 512 | 15 | US-10-106-698-3277 | Sequence 3277, Ap |
| 20 | 51.2 | 3.6 | 1123 | 13 | US-10-425-114-15938 | Sequence 15938, A |
| 21 | 50.6 | 3.5 | 1455 | 15 | US-10-156-761-5889 | Sequence 5889, Ap |
| 22 | 50.6 | 3.5 | 9025608 | 15 | US-10-156-761-1 | Sequence 1, Appl |
| 23 | 50 | 3.5 | 594 | 13 | US-10-142-426-10 | Sequence 10, Appl |
| 24 | 50 | 3.5 | 594 | 15 | US-10-123-155-10 | Sequence 10, Appl |
| 25 | 50 | 3.5 | 594 | 15 | US-10-146-731-10 | Sequence 10, Appl |
| 26 | 50 | 3.5 | 594 | 15 | US-10-140-472-10 | Sequence 10, Appl |
| 27 | 50 | 3.5 | 594 | 15 | US-10-141-761-10 | Sequence 10, Appl |
| 28 | 50 | 3.5 | 594 | 15 | US-10-143-885-10 | Sequence 10, Appl |
| 29 | 50 | 3.5 | 594 | 15 | US-10-158-790-10 | Sequence 10, Appl |
| 30 | 50 | 3.5 | 594 | 16 | US-10-137-871-10 | Sequence 10, Appl |
| 31 | 50 | 3.5 | 594 | 16 | US-10-140-923-10 | Sequence 10, Appl |
| 32 | 50 | 3.5 | 594 | 16 | US-10-141-756-10 | Sequence 10, Appl |
| 33 | 50 | 3.5 | 594 | 16 | US-10-141-759-10 | Sequence 10, Appl |
| 34 | 50 | 3.5 | 594 | 16 | US-10-140-805-10 | Sequence 10, Appl |
| 35 | 50 | 3.5 | 594 | 16 | US-10-140-864-10 | Sequence 10, Appl |
| 36 | 50 | 3.5 | 1362 | 15 | US-10-156-761-7248 | Sequence 7248, Ap |
| 37 | 50 | 3.5 | 9025608 | 15 | US-10-156-761-1 | Sequence 1, Appl |
| 38 | 49.4 | 3.4 | 2336 | 13 | US-10-425-114-15797 | Sequence 15797, A |
| 39 | 48.8 | 3.4 | 919 | 13 | US-10-425-114-22481 | Sequence 22481, A |
| 40 | 48.8 | 3.4 | 2109 | 13 | US-10-425-114-32191 | Sequence 32191, A |
| 41 | 48.6 | 3.4 | 1282 | 13 | US-10-425-114-23114 | Sequence 23114, A |
| 42 | 48.6 | 3.4 | 1339 | 13 | US-10-425-114-33879 | Sequence 33879, A |
| 43 | 48.6 | 3.4 | 1352 | 13 | US-10-425-114-13474 | Sequence 13474, A |
| 44 | 48.4 | 3.4 | 511 | 13 | US-10-424-599-56520 | Sequence 56520, A |
| 45 | 48 | 3.3 | 234 | 10 | US-09-814-353-4810 | Sequence 4810, Ap |

ALIGNMENTS

RESULT 1
US-10-264-237-1097
; Sequence 1097, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAI31P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1097
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-237-1097

Query Match 99.6%; Score 1435.6; DB 16; Length 1474;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1435; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AAAGTAACGGCTACAGACAGTGTGAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTGCG 60
Db 27 AAAGTAACGGCTACAGACAGTGTGAGAAATAGTTTCGCTCGCGGCTAGAAAACTCTGTGCG 86
Qy 61 GTACCAACCCAGAGCGTGTGAGAGCGCCAGCTCCAGCTTCCTTAACGGAGAGTGA 120
Db 87 GTACCAACCCAGAGCGTGTGAGAGCGCCAGCTTCCTTAACGGAGAGTGA 146
Qy 121 GGACTCAGACTTCACAGCCCACTCGGTCGCCAGCTTGTACGCAAGAGACGCCAAGGAC 180
Db 147 GGACTCAGACTTCACAGCCCACTCGGTCGCCAGCTTGTACGCAAGAGACGCCAAGGAC 206

Db 1067 GTTGAAGAGTCCAGACTTAA--TATACACTCTCTGTGAAATAACTGACTCGGA 1123
QY 1239 CTCCTTCCCGCCTTGGACATCCAGCGCGGGAAGCAGTGCCTCGCGCAGGCTG-GGCCA 1297
Db 1124 TTTCTACCCCTCTTTGGGACCCCATAGACTTGAAGACTGTGTTAAGAGCGGTGCCAGGA 1183
QY 1298 GGAGCTCCAGGAGGCACTGAGCGCTGCTGGCGGAGGCTCGGACATCCGACGCA 1357
Db 1184 GCATAGCTTAGCTGAGAACTGTTGCTCCCTGGGAGACAGGAGGCGACCCATGTG 1243
QY 1358 CC-----AGGGAAGTCTCTCGGCGCATCTGTAATAAACCTTTTCTTTTGT 1408
Db 1244 CTTTGTCTAGAGATGGAATTTCTCAGAGAACTGTAAATAAACTTTTGTGTTCTTT 1303
QY 1409 TTTTAAAAAAA 1421
Db 1304 TTTTCTTCAAAA 1316

RESULT 5
US-10-187-657-5
; Sequence 5, Application US/10187657
; Publication No. US20030068311A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy K.W.
; APPLICANT: Baumh, Maria R.
; APPLICANT: Azimzal, Yalda
; TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
; FILE REFERENCE: PV-0009 CLP
; CURRENT APPLICATION NUMBER: US/10/197,657
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: PCT/US00/07817
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/139,565
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030068311A1 3221661R6
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 455, 480, 483
; OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-5

Query Match 31.3%; Score 451.6; DB 15; Length 522;
Best Local Similarity 96.7%; Pred. No. 5e-117;
Matches 502; Conservative 0; Mismatches 12; Indels 5; Gaps 4;

QY 1 AAGTAACGGCTACGACAGTGAAGTAAGTTTCGCTCGCGGCTAGAAAACTCTGTGC 60
Db 9 AAGTAACGGCTACGACAGTGAAGTAAGTTTCGCTCGCGGCTAGAAAACTCTGTGC 68
QY 61 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTTAAAGGAGAGTGA 120
Db 69 GTACCAACCCAGAGCGTTGAGAGCAGCCACCTCCACGCTTCTTTAAAGGAGAGTGA 128
QY 121 GGACTCAGACTTACCAGCCCACTCGCTCCGAGCTTGTACCAAGAGAGAGCCAGGAC 180
Db 129 GGACTCAGACTTACCAGCCCACTCGCTCCGAGCTTGTACCAAGAGAGAGCCAGGAC 188
QY 181 GGCTCTCCGCGTCCAGGAGCCCGCAGCTTGTGCTTGTGCTGCGCGCTGCTGACG 240
Db 189 GGCTCTCCGCGTCCAGGAGCCCGCAGCTTGTGCTTGTGCTGCGCGCTGCTGACG 248
QY 241 ACTCGCGCGGCTGAGCATGACCTGTGGAACGGGTAAGTCTTTTACCCCGAGCC 300

Db 249 ACTCGCGCGGTGAGCATGACCCCTGTGTGAAACGGGTACTGCTTTTACCCCGAGCC 308
QY 301 CGCATGCGCGAGGCTTCAGCGTTCCACTGCTCATGTTATTCTAGTGTGTTTGGTCTTA 360
Db 309 CGCATGCGCGAGTTTCAAGCGTTCCACTGCTCATGTTATTCTAGTGTGTTTGG-TCGA 367
QY 361 GCAGCAAGCTTCCTGCTCATCTTTCGCGGGATCCGTTGGGCACCTCGCGCTGTTGGTTG 420
Db 368 GCAGCAAG--TTCTGCTCATCTTTCGCGGGATCCGTTGGGCACCTCGCGCTGTTGGTTG 425
QY 421 GTGAGAGTTCTTCTCAGTCTGTTTCATAGCGGAGAAATGTCGCTGTGCACTTCAGTGCA 480
Db 426 GTGAGAGTTCTTCTCAGTCTGTTTCATAGGNGCAGAAATGTCGCTGTGCACTTCAGTGCA 484
QY 481 GAATGGTTCTGCGGTACAGTGAAACCAACACATCCTTAC 519
Db 485 GAATGGTTCTGCGGTACATTTG-ACACCAACACATCCTTAC 522

RESULT 6
US-10-094-749-1195/c
; Sequence 1195, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1195
; LENGTH: 2684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-1195

Query Match 27.5%; Score 396; DB 16; Length 2684;
Best Local Similarity 97.6%; Pred. No. 5e-101;
Matches 402; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1007 TCTGGGTCACTGCGCAACCGCGTCTCTGTCCTTCTCGAGGGCGCGTGTGAGTGC 1065
Db 2363 TTTTGATCCCAACCGCCACAGGCGTCTGTCCTTCTCGAGGGCGCGTGTGAGTGC 2304
QY 1067 TCCAGTATGTTGGCGCCAGCGCTCTTGCACCTTCTGAGACCAAGCGCAAGGACTGCA 1126
Db 2303 TCCAGTATGTTGGCGCCAGCGCTCTTGCACCTTCTGAGACCAAGCGCAAGGACTGCA 2244
QY 1127 GCCAGAGAGAGGGGGCTCAGCTCTTATCTCTCGGCGACCTCTGCAAGAGCGCGCTC 1186
Db 2243 GCCAGAGAGAGGGGGCTCAGCTCTTATCTCTCGGCGACCTCTGCAAGAGCGCGCTC 2184

81 CATCATGATCTTTCTGACTGCTGCGCCACGTTTCATCGTCACTGCTGCGG 140
399 CCACTCGCGCTGGTTTGGTTGGTGGAGAGTTCTTCTCAGTCTGTTTCATAGGCGCAGAAAT 458
141 AAGACGAGGCTGTTCTGCTGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 200
459 TGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 518
201 CCGTGTGTGAATTTCACTTCTGAGTGGTCTGCTGCGGCGAGGTGAGTGGTGGTGGTGGTGGTGG 260
519 CAAGGCTTCAAGCGGCGGCGGCTTACAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 578
261 CAGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 320
579 CATTAATATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 638
321 AGTCAACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
639 CAGCAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
381 CGAGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 440
699 GAAGGCT 758
441 GAAGGCT 500
759 CGGCTGTACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
501 TGGCTATACCGCAGTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 995
681 CACATCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740
996 CGGCGCT 1055
741 TGGGCT 800
1056 CGTGTGAGTCTCAGTATGTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115
801 TATGCGGTGGCCACAGGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860
1116 CAAGGA 1121
861 GGATGA 866

RESULT 9

US-10-042-431-54
; Sequence 54, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-6U2
; CURRENT APPLICATION NUMBER: US/10/042.431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063

; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-042-431-54

Query Match 24.6%; Score 354.8; DB 14; Length 1029;
Best Local Similarity 64.5%; Pred. No. 1.5e-89;
Matches 546; Conservative 0; Mismatches 297; Indels 3; Gaps 1;

QY 279 ACTGCTTTTATACCCCGGCGGCGATGCGGAGGCTTACAGCTTCCATGCTCATGCT 338
DB 21 ATTCCCTTCTATGCTGGCCCAAGCCAACTTCCGATGGACACCACTTTGGCCAGCAT 80
QY 339 TATTCTAGTGTGTTTGGCTCTAGCAGCAAGCTTCTCTGCTCATCTTGGCGGATCCGCTGG 398
DB 81 CATCATGATCTTCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 140
QY 399 CCACTCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 458
DB 141 AAGACGAGGCTGTTCTGGCTGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 200
QY 459 TGTGGCTGTGCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 518
DB 201 CTTGGCTGTGATTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 260
QY 519 CAAAGCTTCAAGCGGCGGCTTACAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 578
DB 261 CAAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT 320
QY 579 CATTAATATTAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 638
DB 321 AGTCAACATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380
QY 639 CGAGCAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
DB 381 CGAGGAGTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 440
QY 699 GAAAGGCT 758
DB 441 GAAAGGCT 500
QY 759 CGGCTGTATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
DB 501 TGGCTATACCGCAGTACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560
QY 819 GTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 878
DB 561 ATTCTGCT 620
QY 879 CTTGGCT 935
DB 621 CTACATGCTATTTGGCCACGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680
QY 936 CATCTAGCT 995
DB 681 CACATCACTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740
QY 996 CGGCGCT 1055
DB 741 TGGGCT 800
QY 1056 CGTGTGAGTCTCAGTATGTTGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1115
DB 801 TATGCGGTGGCCACAGGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 860
QY 1116 CAAGGA 1121
DB 861 GGATGA 866

RESULT 10
US-09-759-130B-423
; Sequence 423, Application US/09759130B
; Publication No. US2003002279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kites, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MP100-5350XNIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 423
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-130B-423
Query Match 24.6%; Score 354.8; DB 10; Length 2133;
Best Local Similarity 63.8%; Pred. No. 2.1e-89;
Matches 555; Conservative 0; Mismatches 312; Indels 3; Gaps 1;
QY 255 CAGCATGACCCCTGTGGACGGGTACTGCTTTTACCCCGAGCCCGCGCATGCCGCGAGG 314
DB 66 CAAGATGGCTACTTTGGGACACACATCCCTTCTATGCTGGCCCGCAGCCCACTTCC 125
QY 315 CTTACGGTTCACCTGCTCATGTTTCTAGTGTGTTTGGCTCTAGCAGCAAGCTTCT 374
DB 126 GATGNACACCACTTTGGCCAGCATCATCATGATCTTTCTGACTGCACTGGCCACGTT 185
QY 375 GCTCATCTTGGCGGGATCCGTGGCCATCGCGCTGGTTTGGTTGGTGAGATTTCT 434
DB 186 CGTCATCTCTGCTGGCAATTCGGGGAAGACGAGGCTGTCTGGCTCTGGGGTGGTAC 245
QY 435 CAGTCTGTTTCATAGGCGGAGAAATGTGGCTGTGCATTTCACTGAGAGATGTTCTGGG 494
DB 246 CAGCTTATTCATCGGGGCTGCATCTGCTGTGAATTCAGTTCTGAGTGTCTGTGGG 305
QY 495 TACAGTGAACACCAACACATCTCTACAAAGCCCTTACGCGCAGCGCGGTACAGCCCGGTGT 554

DB 306 CCAGGTCCAGCACCAACACATCATACAGGCCTTCAGTTCTGAGTGGATCAGCGCTGATAT 365
QY 555 CGGTCTGCTGCTGGGCTGGAGGGGCTTATATATTACTCAGAGGACCCCACTGATCA 614
DB 366 TGGGCTGCAGGTCGGGCTGGGTGGAGTCAACATCACTACAGGAGACCCCGGTGAGCA 425
QY 615 GCTGAACGAGACCAATTCATACAAACGAGAGTTTCACTGGCGTCTGAAAAGAGAAATTACGC 674
DB 426 GCTGATGAGACCAATTCATACAAACGAGAGTTTCACTGGCGTCTGGTGAGAACTATGC 485
QY 675 CGCGAGTACCGCAACGCACTGAGAGAGGGGCTGGCGGACCCAGTCTCTACCTGGCGGA 734
DB 486 TGAGGAGTGTGCAAGGCTCTGGAGAAGGGGCTGGCCAGACCCCTGTGTGTGTACCTGTA 545
QY 735 GAAGTTCCACACCGAGTAGCCCTTGGCGCTGTACCAACAGTACCACTGCGCGGAGACACTA 794
DB 546 GAAGTTCCATCCAGAGAGCCCAATGTGGCTTATACCGCCAGTACCGCTGCGCGGACACTA 605
QY 795 CGCTCGCCCAACGCTATGGTGGCGTCTGCTTCTGGCTCTCTTCCAACTGTGCTCTC 854
DB 606 CACTCAGCCATGCTATGGTGGCGCAATCTCTGCTGGCTCTGGCCCAATGTGATGCTCTC 665
QY 855 CAGCGCGCCCGCTCTACGAGGCGCTGGCACTGCTGACCAACCGGAGCCCTTGGCGCTCT 914
DB 666 CATGCTGTCTGCTATATATGGTGGCTACATGCTATTTGGCCACCGGCACTCTTCAGCTGT 725
QY 915 CG---GGGTCTTGGCTTGGCCCTCCATCTCTAGCGTGGCGCTCTGCGCTCTCGCCCTAGG 971
DB 726 GGCTCTGCTCTCTTCTCCATGCGCACATCACTCACTCACTGCTGCTGCTGCTGCTGCT 785
QY 972 CTCCTCGCGCTCAGCACTGAGTACGGCGCGCGCTGCTGGTGCAGCTGCGCAACCGCGCT 1031
DB 786 CGCTTCTGTGTGCTGCTACCTACCATGGCGCTGCTTCTGGATCACATGACCAAGCACT 845
QY 1032 CTTGTGCTCTTCTCTCGAGGGCGCGTGTGAGTCTCCAGTATGTTTCGGGCCAGCGCTCT 1091
DB 846 GCTGTGTGTGCTGCTGGCGCTGCTATGGCGGTGGCCACAGGATGAGCTCAGCGCT 905
QY 1092 TCGCACCTTCTGGACCAAGCGCAAGGA 1121
DB 906 GAAGGCTTTCTTCAACCAAGAGTGTGGAAGA 935
RESULT 11
US-10-042-431-53
; Sequence 53, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-602
; CURRENT APPLICATION NUMBER: US/10/042,431
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-431-53
Query Match 24.6%; Score 354.8; DB 14; Length 2133;
Best Local Similarity 63.8%; Pred. No. 2.1e-89;
Matches 555; Conservative 0; Mismatches 312; Indels 3; Gaps 1;

US-10-187-657-6
SEQUENCE 6, Application US/10187657
Publication No. US20030068311A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy K.W.
APPLICANT: Baughn, Maria R.
APPLICANT: Azimzai, Yalda
TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
FILE REFERENCE: PV-0009 CIP
CURRENT FILING DATE: 2002-07-01
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: US/10/187,657
CURRENT FILING DATE: 2002-07-01
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/139,565
PRIOR FILING DATE: 1999-06-16
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PERL Program
SEQ ID NO 6
LENGTH: 506
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Inocyte ID No. US20030068311A1 SEQ003652D1
FEATURE:
NAME/KEY: unsure
LOCATION: 87, 89, 93, 95, 124, 178, 227, 253, 262, 279, 285, 287, 293-294, 298,
LOCATION: 300, 303, 312, 319-321, 323, 325, 327, 332-333, 336, 343-344, 349, 377, 381,
LOCATION: 405, 409-410, 412, 415, 420, 423, 429, 432, 436-437, 444-445, 448, 452, 454,
LOCATION: 456, 469-471, 479, 483, 485, 492, 494, 499, 501, 504
OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-6
Query Match 22.3%; Score 322; DB 15; Length 506;
Best Local Similarity 86.3%; Pred. No. 28-80;
Matches 353; Conservative 0; Mismatches 54; Indels 2; Gaps 2;
QY 464 CTGTGACCTTCAGTCAGAGATGTTCTGTTGGGTACAGTGAACACCAACACATCTCAAAAG 523
DB 17 CTGTGACCTTCAGT-CAGATGTTCTGTTGGGTACAGTGAACACCAACACATCTCAAAAG 75
QY 524 CCTTCAGCGAGCGCGCTTACAGCCCGTGTGGTCTGTCTGTGGCGCTTGGAGGCAATTA 583
DB 76 CCTTCAGCGAGCGCGCTTACAGCCCGTGTGGTCTGTCTGTGGCGCTTGGAGGCAATTA 135
QY 584 ATATTACATCAGCGAGCGCGCTTACAGCCCGTGTGGTCTGTCTGTGGCGCTTGGAGGCAATTA 643
DB 136 ATATTACATCAGCGAGCGCGCTTACAGCCCGTGTGGTCTGTCTGTGGCGCTTGGAGGCAATTA 195
QY 644 AGTTTACCTTGGCGCTTGAAGAGATTTAGCCCGGAGTACGCGACGCACTGGAGAGAG 703
DB 196 AGTTTACCTTGGCGCTTGAAGAGATTTAGCCCGGAGTACGCGACGCACTGGAGAGAG 255
QY 704 GGCTTCGCGAGCGCGCTTACCTGGCGAGAGAGTTTCAACCGAGTAGCCCTTGGCGGC 763
DB 256 GGCTTCGCGAGCGCGCTTACCTGGCGAGAGAGTTTCAACCGAGTAGCCCTTGGCGGC 315
QY 764 TGTACCAACAGTACCACTTGGCGGAGACATACGCTTGGCGGAGCTATGGGTGGCGTCT 823
DB 316 TGTNNNANNGMAGCCANNNTGCGGACNNNTACNTCGGCCACGCTATGGGTGGCGTCT 375
QY 824 GCTTCTGGCTCC-TCTCCAAAGTGTGCTCTTCAACCGCGCGCGCTCT 871
DB 376 GNTTCTGGCTCTTNTTCAAGGTGCTGTGTTCTNNNANGNGGCGCCGCT 424
RESULT 13
US-10-187-657-3
SEQUENCE 3, Application US/10187657
Publication No. US20030068311A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy K.W.
APPLICANT: Baughn, Maria R.
APPLICANT: Azimzai, Yalda
TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
FILE REFERENCE: PV-0009 CIP
CURRENT FILING DATE: 2002-07-01
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/139,565
PRIOR FILING DATE: 1999-06-16
NUMBER OF SEQ ID NOS: 9

QY 255 CAGCATGACCTGTGAAAGCGGTACTGCTCTTTTATACCCCGAGCCCGGATCCCGCAGG 314
DB 66 CAGATGGCTACTTTGGGACACATCTCCCTTCTATGCTGGCCCGGAGCCACCTTCC 125
QY 315 CTTGAGGCTTCACTGCTCATGTTATCTAGTGTGTTTGGCTCTAGCAGCAAGTCTCT 374
DB 126 GATGGACACCATTTGGCCAGCATCATCATGATCTTTCTGACTGCATGGCCACGTTCA 185
QY 375 GCTCATCTTCCGCGGATCGTGGCCACTCGCTGTTTGTGTTGTGAGAGTCTTCT 434
DB 186 GCTCATCTTCCGCGGATCGTGGCCACTCGCTGTTTGTGTTGTGAGAGTCTTCT 245
QY 435 CAGTCTGTTTCATAGGCGCAGAAATTTGGTGTGCTTCACTTCACTGCAAGATGTTCTGGG 494
DB 246 CAGCTTATTCATCGGCGCTGCAATCTCTGCTGTGAATTTCACTTCACTGCAAGATGTTCTGGG 305
QY 495 TACAGTGAACACCAACATCTTCAAGAGCTTCAAGCGAGCGGCTTACAGCCGCTGT 554
DB 306 CCAAGTTCAGCAACCAACATCATCAAGAGGCTTCAAGTCTTCAAGTGAATCAGCGCTAT 365
QY 555 CGGTCTCTGTTGGCGCTTGGAGGCGCATTAATATTACACTCAGAGGAGCCCGAGTCA 614
DB 366 TGGGCTCAGGCTCGGCTGGGTGGAGTCAACATCACACTCAGAGGAGCCCGTGCAGCA 425
QY 615 GCTGAACGAGACCATTTGACTTACAGAGCAGTTTCACTGCGGCTGAGAGAGATTTACGC 674
DB 426 GCTGAATGAGACCATTAATTAAGAGGAGTTCACCTGCGGCTGGGTGAGAGATTTACGC 485
QY 675 CGCGAGTACGCAACGCACTTGAAGAGGCGCTTCCCGAGCCAGTGTCTTACCTGGCGGA 734
DB 486 TGAGAGTGTGCAAGGCTCTGAGAGAGGCGCTTCCAGACCTGTGTGTACCTAGCTGA 545
QY 735 GAAGTTCAACCGAGTACCTTGGCGCTTACAGAGTACACAGTACACCTGGCGGAGACTA 794
DB 546 GAAGTTCACTCAAGAGGCGCATGTGGCTTATACCGCCAGTACCGCTTGGCGGAGACTA 605
QY 795 CGCTTCGCGCAGCTATGGGTGGGTCTTCTGCTTCTGCTTCTTCCAAAGTGTGCTCTC 854
DB 606 CACTTCAGCCATGTTATGGGTGGGTCTTCTGCTTCTGCTTCTGCTGCTGCTGCTCTC 665
QY 855 CACGCGCGCGCTTCAAGAGGCGCTTGGGCACTGCTGACCAACCGAGGCTTGGCGCTCT 914
DB 666 CATGCTGTGCTGTATATGGTGGCTTACATGCTTATGGCCACGCGCATCTTCCAGCTGT 725
QY 915 CG--GGGTCTTGGCTTGGCTTCCATCTTACAGTGGCGCTTGGCGCTTGGCGCTTGG 971
DB 726 GGCTGTGCTTCTTCTTCCATGGCCACATCACTACCTACCTTCCCTTCCCTGACCTGG 785
QY 972 CTCCTCCGCGCTCACTCAGTACGCGCGCGCTTCTGGGTCTGCTGCGTGGCAACCGCGT 1031
DB 786 CGCTTCTGTGCTGATCTACCTGCGCGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTG 845
QY 1032 CTTGCGCTTCTTCTGAGGCGCGCTGCTGAGTCTCAGTATGTTGGCGGCGCGCTCT 1091
DB 846 GCTGTGTGCTGCTGCGCTTGGCTATGGCGGTGGCCACAGGATGCGAGCTCAGGCT 905
QY 1092 TCGCACCTTCTTGGACCAAGGCGCAAGA 1121
DB 906 GAAGGCTTCTTCAACAGAGTGTGATGA 935
RESULT 12
US-10-187-657-6
SEQUENCE 6, Application US/10187657
Publication No. US20030068311A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy K.W.
APPLICANT: Baughn, Maria R.
APPLICANT: Azimzai, Yalda
TITLE OF INVENTION: TRANSMEMBRANE PROTEIN DIFFERENTIALLY EXPRESSED IN CANCER
FILE REFERENCE: PV-0009 CIP
CURRENT FILING DATE: 2002-07-01
PRIOR FILING DATE: 2000-03-22
PRIOR APPLICATION NUMBER: 60/139,565
PRIOR FILING DATE: 1999-06-16
NUMBER OF SEQ ID NOS: 9

```

? SOFTWARE: PERL Program
? SEQ ID NO 3
? LENGTH: 346
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc feature
? OTHER INFORMATION: incyte ID No. US20030068311a1 1752794f6
? FEATURE:
? NAME/KEY: unsure
? LOCATION: 91, 188, 206, 291, 337
? OTHER INFORMATION: a, t, c, g, or other
US-10-187-657-3

```

| Query Match | 20.8%; | Score 300; | DB 15; | Length 346; |
|-----------------------|-----------------|---|-----------|-------------|
| Best Local Similarity | 94.6%; | Pred. No. 2.8e-74; | | |
| Matches 330; | Conservative 0; | Mismatches 15; | Indels 4; | Gaps 2; |
| QY | 927 | CTTGGCTCCATCTCTAGCGCGCTCTGCCGCTTCGCGCTCCTCGCGCTCAC | 986 | |
| Db | 1 | CTTGGCTCCATCTCTAGCGCGCTCTGCCGCTTCGCGCTCCTCGCGCTCAC | 60 | |
| QY | 987 | CACCTCAGTAGCGGCGCGCTTCTGGGTTCAGCTGGGCAACGGCGTCTGTCCTTTCCT | 1046 | |
| Db | 51 | CACCTCAGTAGCGGCGCGCTTCTGGGTTCAGCTGGGCAACGGCGTCTGTCCTTTCCT | 120 | |
| QY | 1047 | CGGAGGGCGCGTGTGAGTCTCCAGTAGTGTTCGGCCGACGGCTCTTTCGACCCCTTCTGGA | 1106 | |
| Db | 121 | CGGAGGGCGCGTGTGAGTCTCCAGTAGTGTTCGGCCGACGGCTCTTTCGACCCCTTCTGGA | 180 | |
| QY | 1107 | CCAAAGGCCCAAGGACTGTGCAGCGCAGGAGAGAGGGGCTCACCTCTTATCTCTCGGCGA-CC | 1165 | |
| Db | 181 | CCAAAGGCCCAAGGACTGTGCAGCGCAGGAGAGAGGGGCTCACCTCTTATCTCTCGGCGACC | 240 | |
| QY | 1166 | CACCTGCACAGCGCGCTCTCCGACCTTAAATGATATCAGCTAACTAACTGTGAGGG | 1225 | |
| Db | 241 | CACCTGCACAGCGCGCTCTCCGACCTTAAATGATATCAGCTAACTAACTGTGAGGG | 300 | |
| QY | 1226 | GACCAATCTGGACTCTCTCCCGCCTCTGGGACATCGCAGCGCGGGAAG | 1274 | |
| Db | 301 | GACCAATCTGGACTCTT---CCCGCTTGGGAATCGCANTCCGGGAAG | 346 | |

```

RESULT 14
US/10-094-749-1195
/ Sequence 1195, Application US/10094749
/ Publication No. US20030219741A1
/ GENERAL INFORMATION:
/ APPLICANT: ISOGAI, TAKAO
/ APPLICANT: SUGIYAMA, TOMOYASU
/ APPLICANT: OTSUKI, TETSUJI
/ APPLICANT: WATANABE, AI
/ APPLICANT: SATO, HIROYUKI
/ APPLICANT: ISHII, SHIZUKO
/ APPLICANT: YAMAMOTO, JUN-ICHI
/ APPLICANT: ISONO, YUUKO
/ APPLICANT: HIO, YURI
/ APPLICANT: OTSUKA, KAORU
/ APPLICANT: NAGAI, KEIICHI
/ APPLICANT: IRIE, RYOTARO
/ APPLICANT: TAMECHIKA, ICHIRO
/ APPLICANT: SEKI, NACHIKO
/ APPLICANT: YOSHIKAWA, TSUTOMU
/ APPLICANT: OTSUKA, MOTOTYUKI
/ APPLICANT: NAGAHARI, KENJI
/ APPLICANT: MASUHO, YASUHIKO
/ TITLE OF INVENTION: NOVEL FULL-LENGTH
/ FILE REFERENCE: 084335/0160
/ CURRENT APPLICATION NUMBER: US/10/094
/ CURRENT FILING DATE: 2002-03-12
/ PRIOR APPLICATION NUMBER: 60/350,435
/ PRIOR FILING DATE: 2002-01-24
/ PRIOR APPLICATION NUMBER: JP 2001-328

```

```

; PRIOR FILING DATE: 2001-09-14
;
; NUMBER OF SEQ ID NOS: 3381
;
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1195
;
; LENGTH: 2684
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
US-10-094-749-1195

```

| Query Match | 17.1% | Score 245.8 | DB 16 | Length 2684 |
|-----------------------|----------------|---|----------|-------------|
| Best Local Similarity | 67.4% | Pred. No. 1.5e-58 | | |
| Matches 362 | Conservative 0 | Mismatches 172 | Indels 3 | Gaps 1 |
| QY | 588 | TACACTCACAGGACCCCAAGTGCATCAGCTGAAACGAGACCAATTGACTACACAGCAGGTT | 647 | |
| DB | 250 | TGCAATCTCTGGGACCCCTGTCAGCAGCTGAATGAGACCATCAATTACACAGGAGGTT | 309 | |
| QY | 648 | CACCTGGCGCTCTGAAGAGAAATTACGCCGCGAGTACGGGAACCACTGGGAGAGGGGCT | 707 | |
| DB | 310 | CACCTGGCGGCTGGGTGAGAACTATGCTCAGGAGTATGCAAGGCTCTGGGAGAGGGGCT | 369 | |
| QY | 708 | CGCGGACCCAGTCTCTACTCTGGCGGAGAAGTTCAACCGAGTAGCCCTTGCGCCCTGTGA | 767 | |
| DB | 370 | GCAGACCTGTGTGTACTAGCTGAGAAGTTCACTCCAAGAAGCCATGTGGCCCTATA | 429 | |
| QY | 768 | CCACAGTACACCTGGCGGGACACTACGCCCTGGCGACGCTATGGGTGGGTTCTGCTT | 827 | |
| DB | 430 | CGCCAGTACCGCTCGCGGGACACTACACCTCAGCCCATGCTATGGGTGGCAATTCCTCTG | 489 | |
| QY | 828 | CTGCTCTCTTCCAAGCTGCTGCTCTCCACGCGGCCCGCTCTACGAGGCGCTTGGCACT | 887 | |
| DB | 490 | CTGCTCTCTGCCAATGTGATGCTCTCCATGCTCTGTGCTGCTATATGCTGGCTACATGCT | 549 | |
| QY | 888 | GCTGACCAACGGAGCCTTCGCGCTCTTGG--GGGCTTGGCTTGGCTTGGCTCCATCTCTAG | 944 | |
| DB | 550 | ATTGGCCACGGGCAATTTCAAGCTGTGTGCTCTGCTCTCTTCTTCCATGGCCACATCACT | 609 | |
| QY | 945 | CGTGGCGCTCTTGCCCGCTCCGGCTAGGTTCTTCGCGCTCACCACTCAGTACGGCGCCGC | 1004 | |
| DB | 610 | CACCTCACCCCTGTGCCCTGCACCTGGGGGCTTCTGTGCTGCATACTACCAATGGGCGCTC | 669 | |
| QY | 1005 | CTTCTGGATCAAGCTGGCAACCGGCGTCTGTGCGCTCTCTCTCTCGAGGGGCGGTGGTAG | 1064 | |
| DB | 670 | CTTCTGGATCACAATTGACCACAGGACTGTGTGTGTGCTGTGGGCGCTTGGCTATGGCGGT | 729 | |
| QY | 1065 | TCTCAGTATGTTGGGCCCGAGCGCTCTTGCAACCTTCTGACACCAAGGCCCAAGGA | 1121 | |
| DB | 730 | GGCCCAACAGATGACGCTTCAAGGCTGAAGGCTTTCTTCAACACAGAGTGTGATGA | 786 | |

```

RESULT IS
US-10-027-632-91970
; Sequence 91970, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

```

Search completed: June 3, 2004, 13:19:21
Job time : 672 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 3, 2004, 09:02:12 ; Search time 4061 Seconds

(without alignments)
10596.258 Million cell updates/sec

Title: US-09-936-456-1

Perfect score: 1441

Sequence: 1 aaataacggctacagacag.....aaaaaaaaaaaaaaaaaaaaa 1441

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_fam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pug:*

27: em_gss_vri:*

28: gb_gss1:*

29: gb_gss2:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 594.6 | 48.2 | 1594 | 11 AK018569 | AK018569 Mus muscu |
| 2 | 657 | 45.6 | 1309 | 11 AK008816 | AK008816 Mus muscu |
| C 3 | 529.6 | 36.8 | 537 | 9 AI821606 | AI821606 nk08all.x |
| C 4 | 515 | 35.7 | 515 | 9 AI826629 | AI826629 wk35e04.x |

| | | | | | |
|------|-------|------|------|-------------|---------------------|
| C 5 | 510 | 35.4 | 510 | 9 AI304327 | AI304327 q057h06.x |
| C 6 | 489.2 | 33.9 | 505 | 9 AI791844 | AI791844 nk08all.y |
| C 7 | 482.4 | 33.5 | 486 | 9 AI660560 | AI660560 we8b08.x |
| C 8 | 466 | 32.3 | 1014 | 13 BY708714 | BY708714 BY708714 |
| C 9 | 446 | 31.0 | 446 | 9 AI984141 | AI984141 wu21c02.x |
| C 10 | 445.4 | 30.9 | 466 | 9 AA573825 | AA573825 nk08all.s |
| C 11 | 434 | 30.1 | 434 | 9 AI991272 | AI991272 wu41n04.x |
| C 12 | 414 | 28.7 | 414 | 9 AA593860 | AA593860 nm19f03.s |
| C 13 | 409.4 | 28.4 | 411 | 9 AI983793 | AI983793 wu30c09.x |
| C 14 | 405 | 28.1 | 405 | 9 AI274929 | AI274929 ql49c11.x |
| C 15 | 400.2 | 27.8 | 405 | 9 AI346155 | AI346155 qp43f12.x |
| C 16 | 392 | 27.2 | 538 | 9 AI660493 | AI660493 we67h02.x |
| C 17 | 388.4 | 27.0 | 538 | 13 BX517021 | BX517021 BX517021 |
| C 18 | 385 | 26.7 | 573 | 14 CB854121 | CB854121 UI-CF-DU1 |
| C 19 | 381 | 26.4 | 381 | 9 AI281211 | AI281211 qk58e08.x |
| C 20 | 375.8 | 26.1 | 737 | 14 CB305399 | CB305399 UI-CF-EN1 |
| C 21 | 373.6 | 25.9 | 660 | 14 CF766127 | CF766127 CBS004870 |
| C 22 | 372.8 | 25.9 | 646 | 13 BY720747 | BY720747 BY720747 |
| C 23 | 367.4 | 25.5 | 713 | 14 CB171450 | CB171450 JID602700 |
| C 24 | 356.4 | 24.7 | 1032 | 29 AY418362 | AY418362 Homo sapi |
| C 25 | 349.8 | 24.3 | 1026 | 29 AY418364 | AY418364 Mus muscu |
| C 26 | 346 | 24.0 | 483 | 9 AI019726 | AI019726 ua94c11.r |
| C 27 | 335 | 23.2 | 348 | 9 AW050605 | AW050605 w219b11.x |
| C 28 | 333.4 | 23.1 | 690 | 12 BM977010 | BM977010 UI-CF-EN1 |
| C 29 | 333.4 | 23.1 | 896 | 12 BG968477 | BG968477 602835393 |
| C 30 | 324 | 22.5 | 687 | 14 CB850544 | CB850544 UI-CF-EN1 |
| C 31 | 322.4 | 22.4 | 324 | 12 BM987789 | BM987789 UI-H-CO0- |
| C 32 | 308.8 | 21.4 | 871 | 12 B1526485 | B1526485 602923070 |
| C 33 | 304.8 | 21.2 | 919 | 12 B1412602 | B1412602 602990526 |
| C 34 | 296 | 20.5 | 296 | 9 AI821178 | AI821178 nei17a03.y |
| C 35 | 285.2 | 20.5 | 999 | 12 B1410133 | B1410133 602964073 |
| C 36 | 280.6 | 20.2 | 315 | 9 AA618335 | AA618335 ng15g11.s |
| C 37 | 289.4 | 20.1 | 895 | 12 B1413312 | B1413312 602986494 |
| C 38 | 277 | 19.2 | 277 | 9 AI732165 | AI732165 nei17a03.x |
| C 39 | 275 | 19.1 | 608 | 14 CF108325 | CF108325 Shu1tsomi |
| C 40 | 274.4 | 19.0 | 654 | 10 BE914798 | BE914798 601667822 |
| C 41 | 263.2 | 18.3 | 766 | 29 AY418363 | AY418363 Pan trogl |
| C 42 | 260 | 18.0 | 260 | 9 AW009962 | AW009962 ws89h06.x |
| C 43 | 257.2 | 17.8 | 675 | 12 B1154672 | B1154672 602904653 |
| C 44 | 252 | 17.5 | 252 | 9 AI262416 | AI262416 qk38e04.x |
| C 45 | 251.8 | 17.5 | 1040 | 12 B1554670 | B1554670 603236624 |

ALIGNMENTS

| | | | | | |
|------------|--|-------------|------|--------|-----------------|
| RESULT 1 | AK018569 | 1594 bp | mRNA | linear | HTC 20-SEP-2003 |
| LOCUS | AK018569 | | | | |
| DEFINITION | Mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:9030623N16 product:hypothetical protein, full insert | | | | |
| ACCESSION | AK018569 | | | | |
| VERSION | AK018569.1 | GI:12858338 | | | |
| KEYWORDS | HTC; CAP trapper. | | | | |
| SOURCE | Mus musculus (house mouse) | | | | |
| ORGANISM | Mus musculus | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | |
| AUTHORS | Carninci, P. and Hayashizaki, Y. | | | | |
| TITLE | High-efficiency full-length cDNA cloning | | | | |
| JOURNAL | Meth. Enzymol. 303, 19-44 (1999) | | | | |
| MEDLINE | 99279253 | | | | |
| PUBMED | 10349636 | | | | |
| REFERENCE | | | | | |
| AUTHORS | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | | | | |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | | | | |
| JOURNAL | Genome Res. 10 (10), 1617-1630 (2000) | | | | |
| MEDLINE | 20499374 | | | | |
| PUBMED | 11042159 | | | | |

```
end: Baum1. Host: Chn1.
Location/Qualifiers
i. .1394
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9030623N16"
/db_xref="MGI:1894813"
/db_xref="taxon:10090"
/clone="9030623N16"
/sex="male"
/tissue type="colon"
```

| | | | | |
|------------|------|----------------|--|------|
| QY | 1002 | CGCCTTCTGGGTCA | CGCTGGCAACCGGCGTCTGTGCTCTTCTCTCGAGGGGCGCTGGT | 1061 |
| DB | 887 | CTCCTTTTGGCTCA | CGCTGGCAACCGGCGTCTGTGCTCTTCTCTCGAGGGGCGCTGGT | 946 |
| QY | 1062 | GAGTCTCAGTATG | TTCGGCCCGAGCGCTCTTTCGCAACCTCTTGGACCAAGCGCCCAAGGA | 1121 |
| DB | 947 | GATTTCTCCACTA | CTACTCGGCCCGAGCGCTCTTCTCTCGAGGGGCGCTGGT | 1006 |
| QY | 1122 | CT--- | GCAGCCAGGAGAGGGGGGCTCACCTCTTATCTCTCGGCGACCCACTGCGCAAGCA | 1178 |
| DB | 1007 | CTGTAGCAACCA | GGCTTAAGGAACTCACCTCTCACCTCAACCAACCGGCAACACGAAACA | 1066 |
| QY | 1179 | GGCGGCTCTCC | CAGACTTAAATATGATCACCACCTTAACCTGTGTGAGGGGACCCCAATCTGGA | 1238 |
| DB | 1067 | GTTTCAAGAGTCC | AGACTTAAAGGAACTCACCTCTCACCTCAACCAACCGGCAACACGAAACA | 1123 |
| QY | 1239 | CTCCTTCCCGCG | CTTGGGACATCGCAGGCGCGGAAAGCAGTGTGCGCCGCGCAGGCGCTG-GGCCA | 1297 |
| DB | 1124 | TTTCTACCCCT | CTTTTGGGACATCGCAGGCGCGGAAAGCAGTGTGCGCCGCGCAGGCGCTG-GGCCA | 1183 |
| QY | 1298 | GGAGAGCTCAG | GAAGGCGCTGAGCGCTCTGTGGCGGAGGCTCTGACATCTCGCAGGCA | 1357 |
| DB | 1184 | GCATAGCTTAG | TCTGGAGAACTGCTTGTCTGCGGACACGAAAGGCGGCGCAGGCGCTG-GGCCA | 1243 |
| QY | 1358 | CC----- | AGGGAAGTCTCTGCGGGCGGATCTGTAATAAACCCTTTTCTTTTCTTTT | 1408 |
| DB | 1244 | CTTTGCTCTAG | AGATGATTTCTCAGGAGACTGTAAATAAACCTTTTCTTTTCTTTT | 1303 |
| QY | 1409 | TTTTTAAAAAAA | 1421 | |
| DB | 1304 | TTTTTCTCAAAA | 1316 | |
| RESULT 2 | | | | |
| AK008816 | | | | |
| LOCUS | | | | |
| DEFINITION | | | | |
| ACCESSION | | | | |
| VERSION | | | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| MEDLINE | | | | |
| | | | | |

11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furumoto, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, Y., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

DIRECT SUBMISSION
TITLE Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]
JOURNAL
REFERENCE
AUTHORS Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

COMMENT CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCACGACTCTTTTTCCTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATTCGTGGATTCAATAAATTAATCCCCCCCCCCC 3'] cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.

Host: SOLR

FEATURES
source Location/Qualifiers
1..1309
 `/organism="Mus musculus"`
 `/mol_type="mRNA"`
 `/strain="CS7BL/6J"`
 `/db_xref="FANTOM_DB:2210403N03"`
 `/db_xref="WGI:1893739"`
 `/db_xref="taxon:10090"`
 `/clone="2210403N03"`
 `/sex="male"`
 `/tissue type="stomach"`
 `/clone_lib="RIKEN full-length enriched mouse cDNA library"`
 `/dev_stage="adult"`
142..1104
 `/note="unnamed protein product; hypothetical protein (evidence: ProCrest, decoder, NCBI CDS Predictor, Longest-ORF putative)"`
 `/codon_start=1`
 `/protein_id="BAB25910.1"`
 `/db_xref="GI:12843239"`

CDS
 `/translation="MTAMDGVLPFPQPRHAASFVPLLIVTLVPLSLAASFLFILPQG IGHWGNFTLVRLLSLFGISAVIAVFHSGDFWFEVNTNTSKAFSPSRQVTFVLGH VGLAGNVLTFRPTQLNQNTIDYNERFTWRNLNDITKEVTHAEKGLFDPPVLYLAEX FTTPSPGCLYHPFHGYAAHLVGLLEDHRQAALHARPLRGFAVHFHRCLEA IACLRFIDFOSAALPLPWMLRRLLALLLTATLCITSLLEGGAIVHTLRTPR`


```
1. .515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2417406"
/sex="male"
/tissue_type="normal pro
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr-2
/note="Organ: prostate;
with a modified polylink
from normal prostate bul
a Not I - oligo(dT) prim
ligated to Eco RI adapt
I and cloned into the Not
modified pRT73 vector.
constructed by Bento Soa
```

| | | |
|------------|--|---|
| 1384 | GTAATTAACCTTTTCTTTCTTTGTTTTT | 1413 |
| 30 | GTAATTAACCTTTTCTTTCTTTGTTTTT | 1 |
| RESULT 6 | AI791844 | 505 bp mRNA linear EST 13-DEC-1999 |
| LOCUS | nk08a11.55 | NCI CGAP Co2 Homo sapiens cDNA clone IMAGE:1012892 5' |
| DEFINITION | | similar to contains Alu repetitive element; mRNA sequence. |
| ACCESSION | AI791844 | |
| VERSION | AI791844.1 | GI:53339486 |
| KEYWORDS | EST. | |
| ORGANISM | Homo sapiens | |
| ORGANISM | Homo sapiens (human) | |
| REFERENCE | | Bumkayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| AUTHORS | | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| TITLE | 1 (Bases 1 to 505) | |
| JOURNAL | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . | |
| COMMENT | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), | |
| | Tumor Gene Index | |
| | Unpublished (1997) | |
| | Other ESTs: nk08a11.s1 | |
| | Contact: Robert Strausberg, Ph.D. | |
| | Email: cgapbs-remail@nci.nih.gov | |
| | Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. | |
| | Emmert-Buck, M.D., Ph.D. | |
| | cDNA Library Preparation: Stratagene, Inc., David B. Krizman, | |
| | Ph.D. | |
| | cDNA Library Arraying: Greg Lennon, Ph.D. | |
| | Cloning Sequencing by: Washington University Genome Sequencing Center | |
| | Clone Distribution: NCI-CGAP clone distribution information can be | |
| | found through the I.M.A.G.E. Consortium/LLNL at: | |
| | www-bio.lnl.gov/bbrp/image/image.html | |
| | This read is a RESEQUENCE of a previously sequenced human clone | |
| | containing the same sequence. For more information, see original entry for original citation | |

```

information
This 5', resequenced clone has no previous 5' data to verify this
new read against
Putative full length read
The vector to vector length is 671
Insert Length: 741 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 418.
Location/Qualifiers
1. 505
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1012892"
/tissue_type="tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI CGAP Co2"
/note="organ: colon; Vector: Bluescript SK-; Site 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dr. Bulk colon villous adenoma. 5' adaptor sequence:
5' GAATTCGGCAGCAG 3' 3' adaptor sequence: 5'
CTCGAGTTTCTTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."

ORIGIN

Query Match 33.9%; Score 489.2; DB 9; Length 505;
Best Local Similarity 99.2%; Pred. No. 8.3e-57;
Matches 502; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 736 AAGTTTCACACCGAGTAGCCCTTGGCGCTGTATACCCAGTACACCTGGCGGGACACTAC 795
Db 1 AAGTTTCACACCGAGTAGCCCTTGGCGCTGTATACCCAGTACACCTGGCGGGACACTAC 60
|||
QY 796 GCCTGGGCCACGCTATGGTGGCGCTTCGCTTCTGGCTCTCTCCAACTGTGCTCTCC 855
Db 61 GCCTGGGCCACGCTATGGTGGCGCTTCGCTTCTGGCTCTCTCCAACTGTGCTCTCC 120
|||

```


primed with a primer [5', GAGGAGAGAGCGCGCGACACTGGTCTTTTCTTTTCTTTCVN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATCTCGAGTATTAAATTAATCCCCCCCCTCC 3']. cDNA was cleaved with XhoI and SstI.

ORIGIN

Query Match 32.3%; Score 466; DB 13; Length 1014;
Best Local Similarity 80.6%; Pred. No. 7.2e-54;
Matches 569; Conservative 0; Mismatches 135; Indels 2

[illegible]

| | | |
|-----|--|-----|
| 282 | GCCTTTTACCCCCAGCCCGGCATGCGCAGGGTTTCAGGGTTCCACTGCTCATCGTTAT | 341 |
| QY | | |

Db 165 ACCCTTTTACCCCCAGCCAGGCATGCGCGAGCTTCAGCGTACCGCTGCTCATTTGTTAT 224

| Qy | 342 | 225 | 284 |
|----|----------|----------|----------|
| | TCTAGTGT | TCTAGTGT | TCTAGTGT |
| | TTTGGTCT | TTTGGTCT | TTTGGTCT |
| | ATAGCAG | ATAGCAG | ATAGCAG |
| | CAAGCTT | CAAGCTT | CAAGCTT |
| | CCTGCTC | CCTGCTC | CCTGCTC |
| | ATCTTGC | ATCTTGC | ATCTTGC |
| | CGCGGG | CGCGGG | CGCGGG |
| | ATCCGTC | ATCCGTC | ATCCGTC |
| | GCACCA | GCACCA | GCACCA |
| | 401 | 284 | 284 |

402 CTCCGCTGTTTTGGTTGGTTGAGATTCTTCTCAGTCTCTTCATAGCGCGAGAAATTGT 461

| | | | |
|----|-----|--|-----|
| Db | 285 | CTCGCGCTGGTCTCTGGTTGGTGAGAGTCTCTTTAGCCTGTTCATAGGTGCGGAAATTGT | 344 |
| | 452 | AGGCGGCGGCTGTCGTCGACATATCGCTCGCGGCTCAGTCGACATACACCTATCTATCAG | 521 |

| | | | |
|----|-----|--|-----|
| Qy | 462 | GGCTGTGCACCTTCACTGCGAGAATGGTTGTCGGTGTAAGTGGACGCCAACCATCTCTACA | 521 |
| | | | |
| Db | 345 | GGCTGTGCACCTTCACTGTCGAACACTGGTTTGTGGAGGAAGTGTGGACGCNACCATCTCTACA | 404 |

QY 522 AGCCTTCAGCGCAGCGCGCGCTTACAGCCCGTGCCGTCCTCTCGTGGGCGCTGGAGGGCAT 581

| | | | |
|----|-----|---|-----|
| Db | 405 | AGCCTTCAGTCCATCCCGTGTCAAGTCCATGTCGCTCGACGTGGGGCGTGGCGGCGT | 464 |
| Ov | 582 | TAAATATTACATCCACAGGGAACCCACAGTGAATCAGTGCACGACACATTCGACTACACGA | 641 |

| | | | |
|----|-----|---|-----|
| QY | 582 | TAAATTACACTCAGGGGACCCGAGTGCATCAGCTGAACAGAGACCATTCAGCTACAACGA | 641 |
| DB | 465 | TAAACATTACACTCCGAGGAACACCGAGGACGAGCTGAACGAGACCATTCAGCTACAATGA | 524 |

642 GCAGTTCACCTGGCGTCTGAAAGAGAAATTACGCCGCGGAGTACGCGAACGCACTGGAGAA 701

| | | | | | | | | | | |
|----|-----|----------|----------|---------|---------|---------|--------|--------|----------|-------------|
| Db | 525 | GCCTTTCA | CTTGGGCT | CTGACGA | ACTACAC | CAAGGAG | TATGTC | CAATGC | CTTTGAGA | 584 |
| QV | 702 | GGGGTC | GCGGAC | CCCGAT | GTCTACT | GTGGCG | GGAAGT | TTCAC | CCGAGT | AGCCCTTGCGG |

| | | |
|-----|--|-----|
| QY | GGGGTGGGAGACAGAGCTCTACCTGGGAGAGAGTTCACACCGAGTACCCCTTCGGG | 781 |
| | | |
| 702 | GGGGTGGGAGACAGAGCTCTACCTGGGAGAGAGTTCACACCGAGTACCCCTTCGGG | |
| | | |
| Db | GGGGTGGGAGACAGAGCTCTACCTGGGAGAGAGTTCACACCGAGTACCCCTTCGGG | 644 |
| | | |
| 585 | GGGGTGGGAGACAGAGCTCTACCTGGGAGAGAGTTCACACCGAGTACCCCTTCGGG | |
| | | |

762 CCTGTACCAACAGTACCACTGGCGGACACTACGCTCGGCCACGCT-ATGGGTGGCGT 820

| | | |
|-----|--|-----|
| 645 | GCCTGTACCAACAATATCACCTGGCGCGTCACTATATGCGCAGCGACACTGTTGAGGTGCAT | 704 |
| 821 | TCTGCTCTGGGCTCCTCTCCAAACGTGTCTGCTCTCCACGCCGGCCCCCTCTACGAGAGGCC | 880 |

Db

705 TCTGCTTCTGGATCATCGACCAATGGCTGCTCTCCATGCGCGCACACCTCTACGGAGGCC 764

| | | | | | |
|-----|------|-----------------|------------|-----------------------|-----|
| | 881. | TGGCACTGCTGACCA | CGGAGCCATT | CGCGCTCTTCGGGGTCTTGCG | 926 |
| | | | | | |
| bQy | | | | | |
| | 745 | GTTGTATTAATCAGC | CAGCGGCCTA | CTTTCGCGCTCATTCCTG | 808 |
| | | | | | |
| bB | | | | | |

Db 765 TGGCTTTGCTCACCACC-GTGCCTTCGGGCTCTTCGGTGCTTCCC 809

| RESULT 9 | linear | EST 27-OCT-1999 |
|------------|----------|-----------------|
| AI984141/c | 446 bp | |
| LOCUS | AI984141 | |
| | 446 bp | |
| | mRNA | |

LOCUS: 132622.1
DEFINITION: Homo sapiens CDNA clone
IMAGE: 2520574 3' mRNA sequence.
ACCESSION: AT984141

ACCESSION
AI984141
VERSION AI984141.1
KEYWORDS EST. GI:5811360

| SOURCE | ORGANISM |
|----------------------|----------|
| Homo sapiens (human) | |
| Homo sapiens | |
| Homo sapiens | |

FEATURES
source

```

further details.
Location/Qualifiers
1. .1014
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="2210403N03"
/sex="male"
/tissue_type="stomach"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male
stomach"
/notes="Site 1: XhoI; Site 2: SseI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in RIKEN
contributed to prepare mouse tissues. 1st strand cDNA was

```


| | | | |
|----|------|---|------|
| Qy | 1101 | TCTGGACCAAGCGCCAAAGGACTGCAAGCAGGAGAGAGGGGGCTCACCTCTTATCCTCGG | 1160 |
| Db | 314 | TCTGGACCAAGCGCCAAAGGNCCTGCAGCCAGGAGAGAGGGGGCTCACCTCTTATCCTCGG | 255 |
| Qy | 1161 | CGACCCACTGCAACAAGCAGGCGCGCTCCAGACTTAAATGATCACCACCTAACTGTG | 1220 |
| Db | 254 | CGACCCACTGCAACAAGCAGGCGCGCTCCAGACTTAAATGATCACCACCTAACTGTG | 195 |
| Qy | 1221 | AGGGGGACCAATCTGGACTCTCTCCCGCGCTTGGGACATCGCAGGCGGGGAACAGTGC | 1280 |

| | | | |
|----|------|---|------|
| QY | 1281 | CGCCCRAGGCTGGGCAGGAGAGCTCCAGGAAGGCACTGAGCGCTGCTGCGCGAGGCC | 1340 |
| Db | 134 | CGCCCRAGGCTGGGCAGGAGAGCTCCAGGAAGGCACTGAGCGCTGCTGCGCGAGGCC | 75 |
| QY | 1341 | TCGACATCCGACGACACACGAGGAAGTCTCTCGGGCGATCTCTAAATAAACCTTTTT | 1400 |
| Db | 74 | TCGACATCCGACGACACACGAGGAAGTCTCTCGGGCGATCTCTAAATAAACCTTTTT | 15 |
| QY | 1401 | TCTTTTGTTTTTA 1414 | |
| Db | 14 | TCTTTTGTTTTTA 1 | |

RESULT 12
 AS593860/c
 LOCUS
 nhl9f03.s1 NCI_CGAP_Col2 Homo sapiens cDNA clone IMAGE:1084349 3',
 DEFINITION
 mRNA sequence.

```

AA593860.1  GI:2408538
VERSION
EST.
KEYWORDS
SOURCE  Homo sapiens (human)
ORGANISM
REFERENCE  Homo sapiens
AUTHORS  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE  1  (bases 1 to 414)
        NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
        Tumor Gene Index
JOURNAL  Unpublished (1997)
COMMENT  Contact: Robert Strausberg, Ph.D.
        Email: cgapbs-re@mail.nih.gov
        Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
        Emmert-Buck, M.D., Ph.D.
        cDNA Library Preparation: Stratagene, Inc.
        cDNA Library Arrayed by: Greg Lennon, Ph.D.
        DNA Sequencing by: Washington University Genome Sequencing Center
        Clone distribution: NCI-CCAP clone distribution information can be
        found through the I.M.A.G.E. Consortium/LLNL at:
        www-bio.llnl.gov/bxrp/image/image.html
        Insert Length: 1204 Std Error: 0.00
        Seq primer: -40ml3 fwd. ET from Amersham
        High quality sequence stop: 406.
        Location/Qualifiers
            1..414
               /organism="Homo sapiens"
               /mol_type="mRNA"
SOURCE
FEATURES

```

```

/ab_xref="taxon:9606"
/clone="IMAGE:1084349"
/sex="mixed"
/tissue_type="colon tumor"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="NCI-CCAP_Col2"
/note="Organ: colon; Vector: Bluescript SK-; Site 1:
Scor1; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Pooled colon tumors. 5' adaptor sequence: 5'
GAATTCGCACGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTTTT 3' Average insert size: 1.2 kb."
ORIGIN

```

Query Match 28.7%; Score 414; DB 9; Length 414;
 Best Local Similarity 100.0%; Pred. No. 1.3e-46;
 Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 999 CGCGGCTTCTGGTACAGCTGGCAACCGGCTCTGCTGCTCTCTGCGAGGGCCCGT 1058
 DB 414 CGCGGCTTCTGGTACAGCTGGCAACCGGCTCTGCTGCTCTCTGCGAGGGCCCGT 355

QY 1059 GGTGAGTCTCCAGTATGTTGGCGCCAGCGCTCTTGGCACCTTCTGGACCAAGGCCAA 1118
 DB 354 GGTGAGTCTCCAGTATGTTGGCGCCAGCGCTCTTGGCACCTTCTGGACCAAGGCCAA 295

QY 1119 GACTGAGCAGCAGAGAGAGGGGCTCAGCTTATCTTATCTCTGGGACCCATGCAAGCA 1178
 DB 294 GACTGAGCAGCAGAGAGAGGGGCTCAGCTTATCTTATCTCTGGGACCCATGCAAGCA 235

QY 1179 GCGCGCTCTCCAGACTTAAATATGATCAGCACTTAACTCTGAGGGGGACCCATCTGGA 1238
 DB 234 GCGCGCTCTCCAGACTTAAATATGATCAGCACTTAACTCTGAGGGGGACCCATCTGGA 175

QY 1239 CTCCTTCCCGCTTGGGAATGCGAGCGCGGAGAGAGTCCCGCCAGGCTTGGGCCAG 1298
 DB 174 CTCCTTCCCGCTTGGGAATGCGAGCGCGGAGAGAGTCCCGCCAGGCTTGGGCCAG 115

QY 1299 GAGAGCTCCAGGAGGCACTGAGCGCTGCTGGCGGAGGCTCGGACATCCCGAGGCAC 1358
 DB 114 GAGAGCTCCAGGAGGCACTGAGCGCTGCTGGCGGAGGCTCGGACATCCCGAGGCAC 55

QY 1359 CAGGGAAGTCTCTCGGGCGATCTGTAATAAACCCTTTTCTTTTGTGTTTT 1412
 DB 54 CAGGGAAGTCTCTCGGGCGATCTGTAATAAACCCTTTTCTTTTGTGTTTT 1

RESULT 13
 AI983793/c
 LOCUS
 DEFINITION wu20c09.x1 Soares Dieckgraeft colon_NHCD Homo sapiens cDNA clone IMAGE:2520592 3', mRNA sequence.

ACCESSION AI983793
 VERSION AI983793
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 411)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 400.
 Location/Qualifiers
 1..411
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2520592"
 /tissue_type="colonic mucosa from 3 patients with Crohn's disease"
 /lab_hosts="DH10B (phage-resistant)"
 /clone_lib="Soares Dieckgraeft colon_NHCD"
 /note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site: 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I oligo (dT) primer [5' TGTACCAATCTGAAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library

went through one round of normalization. Tissue samples provided by Dr. Brian Dieckgraeft (Washington University, dieck@im.wustl.edu); colonic mucosa represents a range of disease involvement from moderate to severe Crohn's disease; samples include both perforating (fistulas) and non-perforating samples. Library constructed by Bento Soares and M. Fatima Donaldo.

ORIGIN

Query Match 28.4%; Score 409.4; DB 9; Length 411;
 Best Local Similarity 99.8%; Pred. No. 5.6e-46;
 Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1008 CTGGTACAGCTGGCAACCGGCTCTGCTGCTCTCTGCGAGGGCCCGTGGTGTCT 1067
 DB 411 CTGGTACAGCTGGCAACCGGCTCTGCTGCTCTCTGCGAGGGCCCGTGGTGTCT 352

QY 1068 CAGATATGTTGGCGCCAGCGCTCTTGGCACCTTCTGGACCAAGGCCAAGGACTGCGAG 1127
 DB 351 CAGATATGTTGGCGCCAGCGCTCTTGGCACCTTCTGGACCAAGGCCAAGGACTGCGAG 292

QY 1128 CAGAGAGAGGGGCTCAGCTTATCTCTGGGACCCCACTGCAACAGCGCCCTCT 1187
 DB 291 CAGAGAGAGGGGCTCAGCTTATCTCTGGGACCCCACTGCAACAGCGCCCTCT 232

QY 1188 CCAGACTTAAATATGATCAGCACTTAACTCTGAGGGGGACCCATCTGACTCTCTCC 1247
 DB 231 CCAGACTTAAATATGATCAGCACTTAACTCTGAGGGGGACCCATCTGACTCTCTCC 172

QY 1248 CGCTTGGGACATGCGAGCGCGGAGAGAGTCCCGCCAGGCTTGGGCCAGGAGCTCC 1307
 DB 171 CGCTTGGGACATGCGAGCGCGGAGAGAGTCCCGCCAGGCTTGGGCCAGGAGCTCC 112

QY 1308 AGGAGGGCACTGAGCGCTGCTGGCGGAGGCTCGGACATCCGAGGCACACGAGGAAAG 1367
 DB 111 AGGAGGGCACTGAGCGCTGCTGGCGGAGGCTCGGACATCCGAGGCACACGAGGAAAG 52

QY 1368 TCTCTCGGGCGATCTGTAATAAACCCTTTTCTTTTGTGTTTTTAAAAA 1418
 DB 51 TCTCTCGGGCGATCTGTAATAAACCCTTTTCTTTTGTGTTTTTAAAAA 1

RESULT 14
 AI274929/c
 LOCUS
 DEFINITION ql49c11.x1 NCI CGAP Co8 Homo sapiens cDNA clone IMAGE:1875668 3', similar to contains TAR1.t3 TAR1 repetitive element ;, mRNA sequence.

ACCESSION AI274929
 VERSION AI274929.1
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 405)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/hbrp/image/image.html
 Insert length: 1458 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 395.



Creation date: 07-08-2004
Indexing Officer: NGOLSON - NA-WAL GOLSON
Team: OIPEScanning
Dossier: 09641802

Legal Date: 07-07-2004

| No. | Docode | Number of pages |
|-----|--------|-----------------|
| 1 | SRFW | 1 |
| 2 | SRNT | 266 |

Total number of pages: 267

Remarks:

Order of re-scan issued on